

Verzerrt, Verfehlt, Verfälscht?
oder
Kann man Führungsstärke intuitiv
erkennen?

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2008
vorgelegt
von

Diplom-Psychologin Mareike Hoffmann
aus
Köln, Nordrhein-Westfalen

Referent: Prof. Dr. Detlef Fetchenhauer

Korreferent: Prof. Harald Andreas Euler, Ph.D.

Tag der Promotion: 06.02.2009

Dank

Um eine solche Arbeit zu schreiben, nehme man eine Prise Vergnügen an Wissenschaft und Forschung, einen Spritzer Frustrationstoleranz und eine Hand voll Hartnäckigkeit. Die wichtigste und unverzichtbarste Zutat jedoch sind viele, viele Menschen, die ermutigen, kritisieren, unterstützen, anregen, fördern und fordern. Bei diesen Menschen möchte ich mich an dieser Stelle bedanken.

Zunächst bedanke ich mich bei dem Betreuer meiner Dissertation, Herrn Prof. Dr. Detlef Fetchenhauer für seine Visionen und seine fruchtbaren und konstruktiven Kommentare und Gedanken. Auch meinem Zweitgutachter Herrn Prof. Harald A. Euler gilt mein Dank für die vielen wissenschaftlich anregenden Gespräche und die Einblicke in andere Kulturen und Sitten.

Meinen geschätzten Kollegen Dr. Alexandra Haferkamp, Ingrid Kampkötter, Sebastian Lotz, Julia Sauerbrey, Thomas Schlösser und Julia Pradel verdanke ich den Mut und das Durchhaltevermögen, dieses Projekt anzugehen und fertig zu stellen. Sie alle standen mir bei statistischen, nervlichen und diversen anderen kleineren und größeren Problemen mit Rat und Tat und einer pragmatischen Sichtweise zur Seite. Meinen derzeitigen und ehemaligen Diplomanden gebührt ein spezieller Dank für ihre außergewöhnliche Unterstützung und Einsatzbereitschaft bei der Datenerhebung.

Zuletzt und ganz besonders herzlich danke ich meiner Familie und meinen Freunden für Ihr Verständnis und Ihren Beistand. Simone danke ich für die Erfindung der Balkontherapie und das wiederholte Zurechtrücken meiner Perspektive. Sybille, Caro und Nanni: danke für Eure Hilfe in letzter Minute. Vor allem meinem Freund André danke ich für seinen emotionalen Rückhalt und sein Durchhaltevermögen.

Inhaltsverzeichnis

Abbildungsverzeichnis	IV
Tabellenverzeichnis	IV
1. Einleitung	1
2. Führungsstärke – Eine Annäherung	6
3. Personenwahrnehmung aus kognitiv-sozialpsychologischer Perspektive	9
3.1. Kerngedanken der sozialen Kognitionsforschung	9
3.1.1. <i>Ökonomisierung von Denkprozessen: Heuristiken.....</i>	<i>11</i>
3.1.2. <i>Ökonomisierung von Wahrnehmungsprozessen: Stereotype</i>	<i>13</i>
3.2. Das Geschlechtsstereotyp oder „think manager – think male“	14
3.3. Das Attraktivitätsstereotyp oder „what is beautiful is good“	19
3.4. Das „Lack-of-fit“- Modell.....	22
3.5. Schlussfolgerungen für die Wahrnehmung von Führungsstärke	25
4. Personenwahrnehmung aus evolutionspsychologischer Perspektive	28
4.1. Erkennen von Führungsstärke aus evolutionspsychologischer Perspektive	30
4.2. Empirische Überprüfung: das „Thin Slices“ – Paradigma ..	35
4.2.1. <i>Kernannahmen des „Thin slices“- Paradigmas.....</i>	<i>38</i>
4.2.2. <i>Indikatoren für Genauigkeit</i>	<i>39</i>
4.2.3. <i>Kritische Anmerkungen und Bewertung des „Thin slices“- Paradigmas.....</i>	<i>41</i>
4.3. Eine Integration: schnelle und sparsame Heuristiken	45

4.3.1.	<i>Die Hypothese des „Körnchen Wahrheit“</i>	52
4.4.	Der „sexual attribution bias“	55
5.	Integrative Betrachtung und Schlussfolgerungen für die vorliegende Studie	59
6.	Empirische Überprüfung	63
6.1.	Überblick über die empirischen Studien	63
6.2.	Zur Messung von Führungsstärke	65
6.3.	Studie 1	69
6.3.1.	<i>Methode</i>	70
6.3.2.	<i>Ergebnisse</i>	72
6.3.3.	<i>Diskussion</i>	78
6.4.	Studie 2	80
6.4.1.	<i>Ein Referenzmaß für Führungsstärke: die Verhaltensbeobachtung</i>	81
6.4.2.	<i>Methode</i>	83
6.4.3.	<i>Ergebnisse</i>	90
6.4.4.	<i>Diskussion</i>	96
6.5.	Studie 3	99
6.5.1.	<i>Methode</i>	101
6.5.2.	<i>Ergebnisse</i>	104
6.5.3.	<i>Diskussion</i>	112
6.6.	Studie 4	117
6.6.1.	<i>Ein Referenzmaß für Führungsstärke II: das 360°-Feedback</i>	118
6.6.2.	<i>Methode</i>	121
6.6.3.	<i>Ergebnisse</i>	124
6.6.4.	<i>Diskussion</i>	132
7.	Diskussion der empirischen Befunde	137

7.1. Was bisher geschah... Zusammenfassung und Interpretation der Ergebnisse	141
7.1.1. <i>Zur Nutzung von Geschlecht und Attraktivität bei der intuitiven Zuschreibung von Führungsstärke.....</i>	<i>142</i>
7.1.2. <i>Zur Genauigkeit der Wahrnehmung von Führungsstärke .</i>	<i>147</i>
7.1.3. <i>Kritische Betrachtung der verwendeten Referenzkriterien</i>	<i>152</i>
7.2. Fortsetzung folgt... Aufgaben zukünftiger Forschung	157
7.3. Fazit.....	160
8. Literaturverzeichnis.....	161
9. Anhang	176

Abbildungsverzeichnis

Abbildung 1. Streuung der Urteile über Persönlichkeitsstärke durch weibliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der weiblichen Zielpersonen.	76
Abbildung 2. Streuung der Urteile über Persönlichkeitsstärke durch männliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der weiblichen Zielpersonen.	76
Abbildung 3. Streuung der Urteile über Persönlichkeitsstärke durch weibliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der männlichen Zielpersonen.	77
Abbildung 4. Streuung der Urteile über Persönlichkeitsstärke durch männliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der männlichen Zielpersonen.	77

Tabellenverzeichnis

Tabelle 1. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielperson (*: $p < .05$; **: $p < .01$).	78
Tabelle 2: Itemstatistik und Zusammenhänge innerhalb der Skala Führungsverhalten (*: $p < .05$; **: $p < .01$).	89
Tabelle 3. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielperson (*: $p < .05$; **: $p < .01$).	93
Tabelle 4. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Führungsverhalten, Geschlecht und Attraktivität sowie des Interaktionsterms auf die zugeschriebene Persönlichkeitsstärke (*: $p < .05$; **: $p < .01$).	94
Tabelle 5. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie deren Interaktion auf die Skala Führungsverhalten (*: $p < .05$; **: $p < .01$).	95
Tabelle 6. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke bei Laien getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielpersonen (*: $p < .05$; **: $p < .01$).	107
Tabelle 7. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke bei Experten getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielpersonen (*: $p < .05$; **: $p < .01$).	108
Tabelle 8. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Führungsverhalten, Geschlecht und Attraktivität	

sowie des Interaktionsterms auf das Intuitivurteil der Laien (*: $p < .05$; **: $p < .01$).....	110
Tabelle 9. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Führungsverhalten, Geschlecht und Attraktivität sowie des Interaktionsterms auf das Intuitivurteil der Experten (*: $p < .05$; **: $p < .01$).....	110
Tabelle 10. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie deren Interaktion auf die tatsächliche Führungsstärke (*: $p < .05$; **: $p < .01$).....	111
Tabelle 11. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielperson (*: $p < .05$; **: $p < .01$).....	127
Tabelle 12. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen 360°-Feedback, Geschlecht und Attraktivität sowie des Interaktionsterms auf das Intuitivurteil (*: $p < .05$; **: $p < .01$).....	129
Tabelle 13. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie des Interaktionsterms auf das 360°-Feedback (*: $p < .05$; **: $p < .01$).....	130
Tabelle 14. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen tatsächliche berufliche Position, Geschlecht und Attraktivität sowie deren Interaktion auf die vermutete berufliche Position (*: $p < .05$; **: $p < .01$).....	131
Tabelle 15. Regressionskoeffizienten der binär logistischen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie deren Interaktion auf die tatsächliche berufliche Position (*: $p < .05$; **: $p < .01$).....	132

„Afterwards on becoming very intimate with Fitz-Roy, I heard that I had run a very narrow risk of being rejected, on account of the shape of my nose! He was an ardent disciple of Lavater, and was convinced that he could judge a man's character by the outline of his features; and he doubted whether anyone with my nose could possess sufficient energy and determination for the voyage. But I think he was afterwards well-satisfied that my nose had spoken falsely.“

Charles Darwin (nach Barlow, 2005, S.61)

1. Einleitung

Immer, wenn Menschen sich begegnen, wird beurteilt, geurteilt, oftmals auch verurteilt. Erste Eindrücke, spontane Kategorisierungen und intuitive Einschätzungen sind unausweichlich - Menschen können sich nicht einfach vornehmen, an einem bestimmten Tag oder in einer bestimmten Woche keine Urteile über ihre Mitmenschen zu treffen. Dies ist durchaus sinnvoll, denn erst eine adäquate Bewertung der gegebenen Umstände kann ein der Situation angepasstes Verhalten zur Folge haben (Kanning, 1999). Die Betrachtung dieser Umstände fällt in den Gegenstandsbereich der *sozialen Wahrnehmung*.

Unter diesen Begriff werden einerseits die soziale Bedingtheit von Wahrnehmung und andererseits die Wahrnehmung sozialer Tatbestände, wie die Wahrnehmung anderer Personen, gefasst (Gilovich, Keltner & Nisbett, 2006). Beispielsweise ist die Fähigkeit, unaufrichtige von vertrauenswürdigen Menschen zu unterscheiden unerlässlich, um nicht auf die eine oder andere Weise ausgenutzt oder belogen zu werden. Dies ist nur ein Beispiel für die immense Bedeutsamkeit der zuverlässigen Wahrnehmung und Einschätzung von Persönlichkeit, Einstellungen und Motiven im alltäglichen sozialen Miteinander.

Wann immer Menschen sich in einer gegebenen Situation für eine bestimmte Verhaltensweise - und damit gegen andere - entscheiden, beruht diese Entscheidung auf einer Beurteilung. Sie stellt somit einen zentralen Bestandteil der Verhaltenssteuerung dar.

Die Notwendigkeit einer schnellen Beurteilung von Personen und Situationen hat jedoch auch eine andere Seite. Die alltägliche soziale Wirklichkeit ist viel zu komplex, als dass alle Akteure immer alle Details berücksichtigen können (Bierhoff, 2000). So kommt es oftmals zu vorschnellen Beurteilungen von Personen, zu raschen Einschätzungen, die auf wenigen, sofort wahrnehmbaren Merkmalen beruhen.

So wäre Charles Darwins Forschungsreise auf der Beagle beinahe durch die Form seiner Nase verhindert worden. Auch er wurde beurteilt. Der Kapitän der Beagle, Fitz-Roy, zweifelte daran, dass jemand mit einer Nase wie die von Charles Darwin genügend Energie und Zielstrebigkeit für eine solche Reise aufweise. Ein kurzer Blick, wenige gewechselte Worte entschieden über eine Revolution in der Wissenschaft. Der Kapitän der Beagle war - wie viele andere Menschen zu Beginn des 19ten Jahrhunderts - Anhänger der Theorie von Johann Kaspar Lavater.

Lavater ging davon aus, dass die Physiognomie eines Menschen, das heißt die für ihn charakteristischen Gesichtszüge und seine Körperformen, eine sichtbare Manifestation seiner latenten und damit unsichtbaren Qualitäten darstelle (Hartley, 2001). Damit unterschieden sich seine Annahmen nicht wesentlich von anderen physiognomischen Theorien seiner Zeit. Heutzutage gelten Theorien auf Grundlage der Physiognomie wegen ihres rassistischen Kontextes und ihrer Neigung zu unspezifischen Annahmen und unbeweisbaren Schlüssen als überholt (Collins, 1999).

Der Antritt einer Forschungsreise in der heutigen Zeit würde somit sehr wahrscheinlich nicht mehr maßgeblich von der Nasenform des Passagiers abhängen. Könnte man daraus schließen, dass Menschen

nun weniger voreingenommene Urteile über andere Menschen treffen? Schätzen sie heute die Charakteristiken ihrer Mitmenschen zutreffender ein? Würde ein Kapitän des 21ten Jahrhunderts Charles Darwins Forschernatur und Potenzial als Anführer einer wissenschaftlichen Revolution sofort erkennen?

Mit dieser Frage - und den empirischen Antworten - beschäftigt sich die vorliegende Arbeit. Konkret soll es um die Frage gehen, ob Menschen in der Lage sind, die Führungsstärke anderer Menschen schnell und akkurat einzuschätzen.

Zu diesem Zweck wurde eine Reihe von Studien durchgeführt, bei denen die Teilnehmer tonlose Videosequenzen sahen. Auf 20 Sekunden gekürzte Verhaltensausschnitte zeigten die zu beurteilenden Personen jeweils einzeln an einem Tisch sitzend und in die Kamera sprechend. Aufgabe der Beurteiler war es, in der kurzen Pause zwischen den 20-sekündigen Ausschnitten die soeben beobachtete Person auf ihre Führungsstärke hin einzuschätzen. Aufgrund der Kürze der Zeit zwischen den einzelnen Videosequenzen waren die Beurteiler gezwungen, ihr Urteil rasch und intuitiv zu fällen.

Dieses allgemeine Design wurde in den verschiedenen Studien gezielt variiert, um jeweils unterschiedliche Facetten der allgemeinen Fragestellung zu beleuchten. Zum einen variierten die Beurteiler und die zu beurteilenden Personen hinsichtlich Alter, Berufsstand und Erfahrung mit Beurteilungsprozessen. Zum anderen wurden unterschiedliche Vergleichskriterien zur Messung der Wahrnehmungsgenauigkeit herangezogen. Neben der adäquaten Identifikation von Führungsstärke wurde ebenso untersucht, an welchen Merkmalen sich die Beurteiler bei der Zuschreibung von Führungsstärke orientieren. Ein besonderes Augenmerk lag auf der Frage, ob und in welcher Weise das Geschlecht und die Attraktivität der zu beurteilenden Personen einen Einfluss auf die Zuschreibung von Führungsstärke haben.

Zu Fragen nach dem Prozess und der Genauigkeit menschlicher Wahrnehmung können aus den Erkenntnissen der relevanten For-

schungsströme ganz unterschiedliche Vermutungen formuliert werden.

Betrachtet man die Aussagen und Ergebnisse der sozialen Kognitionsforschung und der evolutionären Psychologie, so fällt auf, dass beide zwar auf den gleichen Ausgangsbeobachtungen fußen, aber dennoch zu unterschiedlichen Schlussfolgerungen für die vorliegende Fragestellung kommen. Beide Forschungsströme gehen davon aus, dass menschliche Wahrnehmungs- und Denkprozesse in einer hochgradig komplexen sozialen Umwelt lediglich auf einem Bruchteil der zur Verfügung stehenden Informationen beruhen (Bierhoff, 2000). Doch während die soziale Kognitionsforschung annimmt, dass dies zu verzerrten Wahrnehmungen und fehlgeleiteten Entscheidungen führt (Krueger & Funder, 2004), geht die evolutionäre Psychologie davon aus, dass Menschen die relevantesten und aussagekräftigsten Informationen selektieren und somit zu guten Urteilen und Entscheidungen gelangen (Gigerenzer, Todd & the ABC Research Group, 2001; Haselton & Funder, 2006).

Der grundlegendste Unterschied zwischen diesen beiden Positionen besteht jedoch in Bezug auf ihre forschungsleitenden Fragen. Während sich die soziale Kognitionsforschung vornehmlich mit Wahrnehmungsprozessen beschäftigt, fragt die evolutionäre Psychologie nahezu ausschließlich nach dem Ergebnis menschlicher Wahrnehmung.

Obwohl eine Paradigmen übergreifende Zusammenarbeit sicherlich fruchtbar wäre, bleiben die Gemeinsamkeiten und bisherigen Erkenntnisse des anderen von beiden Seiten bislang weitgehend unbeachtet.

Insgesamt erweitert diese Forschungsreihe den derzeitigen Wissensstand somit um zwei entscheidende Aspekte. Erstens untersucht sie die Genauigkeit der sozialen Wahrnehmungsfähigkeit in einem bisher unerforschten Bereich. Bezogen aber auf die dahinter stehende allgemeine Frage, ob Menschen grundsätzlich in der Lage sind, andere

Menschen genau und zuverlässig einzuschätzen, leistet sie den größten Beitrag, indem sie die Aussagen mehrerer, bisher unabhängig voneinander existierender Forschungsstränge zusammenführt und einander gegenüberstellt.

Seit Lavater und anderen physiognomischen Theorien haben sich die Laien- sowie die Expertenauffassungen bezüglich des Zusammenhangs zwischen konkreten äußerlichen Merkmalen und Charakter stark verändert. Doch die grundlegenden Wesenzüge des Beurteilungsprozesses bleiben weiterhin gegeben. Urteile über andere Menschen und Entscheidungen in einer sozialen Welt beruhen auf einem Bruchteil der existierenden Informationen und stellen lediglich Wahrscheinlichkeitsaussagen dar, die mit einer gewissen Unsicherheitskomponente behaftet sind (Gilovich, Griffin & Kahnemann, 2002). Wie also wirken sich diese beiden Faktoren auf die Validität der menschlichen Wahrnehmung aus?

Können Menschen auch mit wenigen Informationen gute Entscheidungen fällen? Dies würden zumindest evolutionär orientierte Sozialpsychologen aussagen. Oder liegen Menschen mit ihren Wahrscheinlichkeitsurteilen meist falsch, gerade weil sie nur wenige Informationen in ihr Urteil integrieren? Diese gegensätzliche Position würde eher von kognitiven Sozialpsychologen befürwortet. Mit dem Ziel, diesen Fragen einen wichtigen Puzzlestein hinzuzufügen, wurde dieses Forschungsprojekt begonnen.

Vor der Darstellung und der Diskussion der empirischen Ergebnisse werden die relevanten Forschungsrichtungen zur Personwahrnehmung skizziert. Es wird deutlich werden, dass sich aus den vorgestellten Forschungsrichtungen nahezu gegensätzliche Hypothesen für die vorliegende Fragestellung ergeben. Zunächst jedoch soll das Konstrukt Führungskraft als Gegenstand dieser Forschung eingeführt werden.

2. Führungsstärke – Eine Annäherung

“To an extent, leadership is like beauty; it’s hard to define but you know it when you see it.”

Warren G. Bennis (2003, S.XXVI)

Über das Thema Führung wird in der Wissenschaft wie auch in der Gesellschaft rege diskutiert, geschrieben und geforscht (Bligh & Meindl, 2005). Obwohl es dementsprechend eine Vielfalt an Forschungsergebnissen, theoretischen Ansätzen, Konzepten und Literatur gibt, fehlt bislang eine einheitliche und allgemeingültige Definition von Führungsstärke (Yukl, 2002).

Nach Weiss (2003) gehören zu wahrer Führungsstärke folgende Eigenschaften:

- **Energie:** gute Führungspersonen strahlen Optimismus und Bestimmtheit aus. Sie zeigen ihr eigenes Commitment, indem sie hart arbeiten und somit die Geschwindigkeit und die Standards für andere setzen. Dabei halten sie sich selbst an die gleichen Standards, die sie auch von anderen fordern.
- **Ideen:** Führungsstarke Personen sorgen mit ihren Ideen in nahezu allen Situationen für Innovation und Verbesserung und nehmen damit die Risiken auf sich, die viele Führungsentscheidungen mit sich bringen.
- **Erfolg:** Mitarbeiter vertrauen sich einer Führungsperson bereitwillig an, wenn sie an deren Pläne und Fähigkeiten glauben. Die Führungsperson sucht die Verantwortung für eine Aufgabe und erhält Respekt durch bedeutende Leistung.
- **Soziale Empfänglichkeit:** Wahre Führungspersonen vergewissern sich, was ihre Beschäftigten benötigen und denken.
- **Führung:** andere Menschen in einen Plan zu integrieren und diesen beharrlich voranzutreiben, sowie schnelle Reaktionen

bei auftretenden Herausforderungen zeugt von wahrer Führungsstärke.

- Delegationsfähigkeit: effektive Führer bauen durch eine intelligente und konsistente Delegation von Aufgaben eine vertrauensvolle Beziehung zu ihren Mitarbeitern auf. Diese sammeln so Erfahrungen, verstärken das Vertrauen in ihre eigenen Fähigkeiten und arbeiten produktiv und proaktiv.

Lange Zeit dominierten in der Führungsforschung monokausale Ansätze zur Erklärung von Führungsstärke. Nachdem jedoch weder alleine durch konkrete Eigenschaften der Führenden (Lord, De Vader & Allinger, 1986; Mann, 1959; Stogdill, 1948), noch durch spezifische Führungsstile (Seidel, 1978; von Rosenstiel, 2006), noch durch kontextuelle beziehungsweise situative Faktoren (Fiedler, 1967; Fiedler & Macaulay, 1998; Vroom, 2000; Vroom & Yetton, 1973) ein substantieller Anteil von Führungserfolg erklärt werden konnte, stehen heute Modelle im Vordergrund, die mehrere dieser Faktoren integrieren (von Rosenstiel, 2006).

Eines dieser Modelle und gleichsam die derzeit populärste Führungstheorie ist die der charismatischen Führung (Bono & Judge, 2004). Durch diesen im Kern eigenschaftstheoretischen Ansatz erfuhren die Charakteristiken von Führungskräften eine bemerkenswerte Renaissance (Connelly, Gilbert, Zaccaro, Threfall, Marks & Mumford, 2000).

Die Theorie der charismatischen Führung geht auf Weber (1947) zurück, der Führungserfolg vor allem auf die Wahrnehmung der Führungsperson als Person besonderer Qualität zurückführte. Charisma wird in modernen Ansätzen als ein Konstrukt verstanden, welches sich in bestimmten Situationen in ein bestimmtes Verhalten des Führenden übersetzt (von Rosenstiel, 2006). Kernpunkt dieses Ansatzes ist, dass Mitarbeiter das Verhalten ihrer Führungskraft als außergewöhnlich und herausragend wahrnehmen (Crant & Bateman, 2000). Infolgedessen identifizieren sie sich mit ihrer Führungsperson und betrachten sie als Vorbild, sie internalisieren deren Werte und

Ziele und lenken ihre Ressourcen und Energien auf die Ziele der Führungskraft.

Gleichwohl ist die Fokussierung auf Charisma als relevante Führungseigenschaft nicht als Rückfall in einen simplen monokausalen Erklärungsansatz zu verstehen. Vielmehr realisiert sich charismatische Führung erst durch das Verhältnis von Mitarbeiter und Führendem. Dazu gehören sowohl bestimmte Eigenschaften des Führenden wie Extraversion und Verträglichkeit (Bono & Judge, 2004), Selbstvertrauen, prosoziale Durchsetzungskraft, Kreativität, Entscheidungsfreudigkeit, Einfühlungsvermögen, etc. (House & Howell, 1992). Charismatische Führer sind in der Regel selbstbewusste Personen, die über klare Visionen verfügen, unkonventionelle Verhaltensweisen zeigen und sich in Veränderungsprozessen als treibende Kraft beweisen (Shamir, House & Arthur, 1993; Shamir, Zakay & Breinin, 1998).

Gleichzeitig zeigen charismatische Führer bestimmte Verhaltensweisen, wie die individuelle Berücksichtigung der Untergebenen, inspirierende Motivation, ausdrucksstarke Kommunikation, Eingehen persönlicher Risiken, Aussprechen von Vertrauen in die Geführten, intellektuelle Stimulation, etc. (Yukl, 2002).

In Anlehnung an diese integrative Sichtweise wird im Rahmen der vorliegenden Arbeit Führungsstärke im Sinne von charismatischer Führung verstanden.

Dabei stehen auf der einen Seite Eigenschaften, die eine Person mitbringen sollte, beispielsweise Führungsmotivation (Weiss, 2003), interpersonelle Sensitivität (Yukl, 2002) oder Durchsetzungsvermögen (House & Howell, 1992) und andererseits konkret führungsstarkes Verhalten, beispielsweise Motivation und Einbezug aller Gruppenmitglieder oder Ausrichtung der Gruppenmitglieder auf ein Ziel.

Insofern besteht eine zentrale Herausforderung dieser Arbeit darin, die Führungsdisposition sowie das Führungsverhalten zu messen, so dass durch den Vergleich von intuitiv wahrgenommener und tatsäch-

lich gemessener Führungskraft eine Aussage über die Validität der Wahrnehmung von Führungskraft möglich wird. Vorab jedoch wenden wir uns der Frage zu, welche bisherigen Erkenntnisse und Theorien im Bereich der Personenwahrnehmung welche Vermutungen über die vorliegende Fragestellung erlauben.

3. Personenwahrnehmung aus kognitiv-sozialpsychologischer Perspektive

Menschen sind hochgradig soziale Wesen. Daher ist der Umgang mit anderen Menschen, genauso wie die daraus resultierenden Probleme, für sie von enormer Bedeutung (Fischer & Wiswede, 2002). Obwohl ihr alltägliches Verhalten und Denken vielfach von eindrucksvollen Fähigkeiten zeugt, konzentriert sich der sozialpsychologische Mainstream seit Jahrzehnten auf die auf vielerlei Arten aufzeigbaren Verfehlungen menschlichen Denkens und Handelns (Krueger & Funder, 2004).

Die letzten drei Jahrzehnte brachten der klassischen Sozialpsychologie eine Neuorientierung. Der Fokus verlagerte sich von sozialem Verhalten auf die Untersuchung sozialer Wahrnehmungs- und Urteilsprozesse. Dieser Wandel ging einher mit einer stärkeren Betonung kognitiver Prozesse, so dass eine neue Forschungstradition, die soziale Kognitionsforschung oder kognitive Sozialpsychologie entstand.

3.1. Kerngedanken der sozialen Kognitionsforschung

Menschen finden sich täglich in einer Vielzahl unterschiedlicher sozialer Situationen wieder, die jeweils schnelle Bewertungen als Grundlage adäquaten Verhaltens erfordern. Angesichts der begrenzten Wahrnehmungs- und Informationsverarbeitungskapazität des Menschen ist eine schnelle Orientierung nur durch zwei Mechanis-

men möglich: Wahrnehmungsselektion und Ökonomisierung von Urteils- und Entscheidungsprozessen (Kanning, 1999).

Würde jeder noch so schwache Reiz, der in jeder Sekunde auf den Organismus einströmt, bewusst wahrgenommen und verarbeitet werden, so wäre dieser kaum noch handlungs- und überlebensfähig. Wäre der menschliche Denkprozess nicht durch bestimmte Hilfsmittel wie Heuristiken vereinfacht, so müsste der alltägliche Handlungsstrom ständig unterbrochen werden, da das sorgfältige Überdenken und Abwägen aller Sachverhalte einen Großteil der verfügbaren Aufmerksamkeit und Informationsverarbeitungskapazität binden würde (Kanning, 1999). Eine Person, die, noch bevor sie in ihr Auto steigt, zunächst sämtliche Risiken und Handlungsszenarien systematisch durchdenken und abwägen müsste, bevor sie sich für eine Handlung, wie beispielsweise die Fahrt zum Arbeitsplatz, entscheiden kann, würde wahrscheinlich nie auf ihrer Arbeitstelle ankommen.

Bilden sich Menschen somit Eindrücke oder fällen sie Urteile über andere Menschen, so unterliegen diese immer zwei Prozessen: Zum einen wird die Komplexität der eingehenden Informationen drastisch reduziert und zum anderen werden kognitive Hilfsmittel verwendet, die schnelle und effiziente Entscheidungen erlauben.

Betrachten wir zunächst den Prozess der Wahrnehmungsselektion. Nicht alle Informationen, die auf den menschlichen Organismus eintreffen, werden von diesem auch wahrgenommen. Fitz-Roy, der Kapitän des Forschungsschiffs Beagle, nahm wahrscheinlich sehr genau die Nasenform von Charles Darwin wahr, möglicherweise jedoch nicht die Form seiner Hände, selbst wenn sein Blick über diese geschweift sein mag.

Die Wahrnehmungsselektion des Kapitäns war psychologischer Natur, da sie durch Lavaters physiognomische Theorie gesteuert wurde. Doch psychologische Faktoren sind nicht die einzigen Ursachen für die Selektion von Informationen. Grobe Schätzungen gehen davon aus, dass von den circa eine Milliarde Informationseinheiten (*bit*), die

pro Sekunde auf den menschlichen Organismus einströmen, nur etwa 100 Bit bewusst wahrgenommen werden (Kebeck, 1997). Im Prozess der Wahrnehmung werden aufgrund der begrenzten Wahrnehmungskapazität des Menschen wichtige von unwichtigen Informationen getrennt. Welche Informationen dabei in das Bewusstsein dringen, hängt von verschiedenen Faktoren wie Intensität, Kontext, Eindeutigkeit, Gewöhnung und Relativität des äußeren Reizes und vom Aufmerksamkeitsfokus des Wahrnehmenden ab (Hussy, 1986).

Unklar ist bisher, wann diese Informationsreduktion stattfindet. Einerseits könnte sie durch eine Beschränkung der neuronalen Übertragungskapazität entstehen (Broadbent, 1958), andererseits könnte sie - zumindest teilweise - das Ergebnis kognitiver Verarbeitungsprozesse sein (Hussy, 1986; Kanning, 1999).

Fest steht jedoch, dass Eindrücke über andere Menschen nicht auf der Gesamtheit der zur Verfügung stehenden Informationen beruhen.

3.1.1.Ökonomisierung von Denkprozessen: Heuristiken

Neben der Selektion ist die effiziente Verarbeitung von Informationen die zweite essentielle Voraussetzung einer schnellen Orientierung in der sozialen Umwelt. Zur Ökonomisierung von Informationsverarbeitungs- und Denkprozessen werden vornehmlich Heuristiken eingesetzt.

Heuristiken sind Urteilstechiken, die zu einer Reduktion der Komplexität von Urteilsaufgaben beitragen. Dabei handelt es sich um automatisch und unbewusst angewendete Daumenregeln, die ohne große kognitive Anstrengung ein schnelles Urteil erlauben (Gilovich et al., 2002).

Ein Großteil der alltäglichen Entscheidungen und Urteile über andere Menschen wird auf Basis von Wahrscheinlichkeiten gefällt. Zum Beispiel ist die Entscheidung eines Personalverantwortlichen, den einen und nicht den anderen Kandidaten einzustellen, Ausdruck seiner Vermutung darüber, mit welcher Wahrscheinlichkeit der eine

Kandidat - im Vergleich zum anderen - die zu besetzende Position ausfüllen und seine Arbeit gut meistern wird. Eine solche Entscheidung ist eine Entscheidung unter Ungewissheit.

Zum einen fließen nicht alle zur Verfügung stehenden Informationen in den Entscheidungsprozess ein. Zum anderen handelt es sich lediglich um eine Wahrscheinlichkeitssaussage, die immer mit einer Unsicherheitskomponente einhergeht. Ob der gewählte Kandidat tatsächlich eine gute Führungskraft abgibt, hängt von verschiedenen Faktoren ab, beispielsweise von der Passung zwischen Kandidat und seinem Team, der Unternehmenskultur, der Unternehmenssituation (Ostroff & Judge, 2007), usw. Doch keinesfalls ist eine exakte Aussage über den Ausgang dieser Situation möglich.

Die Frage, wie Menschen mit Wahrscheinlichkeiten beziehungsweise mit Entscheidungen unter Ungewissheit umgehen, zieht bis heute in der Forschung große Aufmerksamkeit auf sich. Im Mittelpunkt dieses Interesses steht die Identifizierung von Schwächen des menschlichen Entscheidungs- und Urteilsvermögens. Dies geschieht nicht zuletzt aus der Überlegung heraus, dass gerade durch die Analyse von Urteilsverzerrungen und -fehlern die grundlegenden Strukturen dieser Prozesse deutlich werden (Kunda, 1999; Funder, 1987). Ein weiterer Grund ist die allgemeine Relevanz des Themas. Für Menschen, insbesondere für Entscheidungsträger wie Politiker, Richter, Personalverantwortliche usw., ist es wichtig zu wissen, bei welchen Entscheidungen und Urteilen sie typischerweise welchen Fehlern unterliegen können (Evans, 1984).

Der Kerngedanke des so genannten „*heuristics and biases*“- Ansatzes ist, dass Alltagsentscheidungen unter Ungewissheit eher auf wenigen, stark simplifizierenden Heuristiken basieren statt auf intensiven algorithmischen Denkprozessen (Gilovich et al., 2002). Damit stellt dieser Ansatz einen Kontrapunkt zur klassischen „*rational choice*“-Theorie dar, der Annahme, Menschen würden stets alle zur Verfügung stehenden Informationen in verschiedene Handlungsalternati-

ven integrieren, diese vollkommen rational gegeneinander abwägen und sich dann entscheiden (Becker, 1976; Simon, 1955). Auch „*rational choice*“-Vertreter würden nicht behaupten, Menschen lägen bei ihren Entscheidungen immer richtig. Doch sie würden von unsystematisch auftretenden Fehlern ausgehen, während der „*heuristics and biases*“-Ansatz systematische Verzerrungen annimmt (Gilovich & Griffin, 2002).

Einige systematische Urteilsverzerrungen sind dank extensiver Forschung heute gut dokumentiert. Dazu gehören zum Beispiel der fundamentale Attributionsfehler (Nisbett & Ross, 1980), hypothesengeleitete Wahrnehmung (Bruner & Postman, 1951), Überattribution (Jones & Harris, 1967), die Wahrnehmung illusionärer Korrelationen, Priming, retrospektive Irrtümer (Bierhoff, 2000; Kanning, 1999), der Halo-Effekt und viele mehr. Insgesamt lässt sich festhalten, dass sozial-kognitive Ansätze der Personenwahrnehmung vor allem die Notwendigkeit und die Konsequenzen einer ökonomischen Informationsverarbeitung betonen (Cloutier, Mason & Macrae, 2005). Innerhalb dieser Ansätze besteht die grundsätzliche Annahme, die Integration von mehr Informationen in den Entscheidungsprozess verbesserte die Qualität der Entscheidung beziehungsweise das abschließend gefällte Urteil.

3.1.2.Ökonomisierung von Wahrnehmungsprozessen: Stereotype

Neben dem Einsatz von Heuristiken, stellen Kategorisierungsprozesse eine weitere Möglichkeit zur Ökonomisierung des Wahrnehmungsprozesses dar. In diesem Rahmen kommt der Anwendung von Stereotypen im Bereich der Personenwahrnehmung eine zentrale Bedeutung zu (Wänke, Bless, & Wortberg, 2003).

Stereotype sind mentale Repräsentationen, die Wissen, Überzeugungen und Erwartungen bezüglich sozialer Gruppen und Rollen beinhalten (Kunda, 1999; Pawlik, 2006). Unter Anwendung von Stereotypen orientieren sich Einschätzungen über andere Menschen vor

allem an salienten Merkmalen, beispielsweise Geschlecht, Alter, Hautfarbe, etc. Sie beinhalten, dass Personen nicht als Individuen beurteilt werden, sondern als Teil einer Gruppe, da ihnen die vermeintlichen Charakteristiken der Gruppe oder der sozialen Rolle zugeschrieben werden. Auch Stereotype werden zum „*heuristics and biases*“ – Ansatz gezählt, da sie im Prinzip eine Überschätzung der Basisrate einer Eigenschaft innerhalb einer bestimmten Gruppe darstellen (Locksley, Hepburn & Ortiz, 1982).

Ein verbreitetes Stereotyp ist beispielsweise, Männer seien assertiver als Frauen. Daraus folgt, dass der Anteil durchsetzungsfähiger Männer in der Bevölkerung höher sein sollte als der Anteil assertiver Frauen. Damit ist die subjektive Wahrscheinlichkeit, einen assertiven Mann zu treffen, höher als die, eine durchsetzungsfähige Frau zu treffen (Locksley, Borgida, Brekke & Hepburn, 1980). In diesem Sinne sind stereotype Vorstellungen eine Art Alltags-Heuristik, wobei die Gruppenzugehörigkeit als Hinweis auf die Fähigkeiten und Eigenschaften einer Person herangezogen wird. Sie haben neben dem Nutzen der Vereinfachung und schnellen Orientierung die wesentliche Funktion der Vorhersage von Verhalten (Swazina, Waldherr & Maier, 2004).

Zwei intensiv erforschte, gut dokumentierte und für die Zuschreibung von Führungsstärke besonders relevante Stereotype sind das Geschlechts- und das Attraktivitätsstereotyp.

3.2. Das Geschlechtsstereotyp oder „think manager – think male“

„Women who seek to be equal with men lack ambition.“

Timothy Leary (1920-1996)

Für schnelle Urteile und grundlegende Einordnungen anderer Menschen bietet sich eine Kategorisierung anhand des biologischen Ge-

schlechtes als salientestes Merkmal einer Person besonders an. Auch wenn keine weiteren Informationen über eine Person bekannt sind, ist das Geschlecht sofort wahrnehmbar und somit für die Eindrucksbildung leicht nutzbar. Im Sinne einer schnellen Bewertung und Orientierung in sozialen Situationen wird bei Kontakt mit fremden Menschen aus der Unkenntnis in Bezug auf weitere Gruppenzugehörigkeiten das Geschlecht als Basiskategorie verwendet (Swazina et al., 2004). Diese erste Kategorisierung evoziert allgemeine mentale Assoziationen über Eigenschaften und Erwartungen an das Verhalten und Denken von Männern und Frauen. Diese führen dazu, dass das Individuum stereotypähnlicher wahrgenommen wird als es in Wirklichkeit ist (Eagly & Sczesny, im Druck).

Definitiv sind deskriptive Stereotype von Stereotypen mit präskriptivem Inhalt zu unterscheiden. Das deskriptive Geschlechtsstereotyp beschreibt die Vorstellung, wie Männer und Frauen typischerweise sind, während das präskriptive Geschlechtsstereotyp beinhaltet, wie diese idealerweise sein sollten (Deaux & LaFrance, 1998). Vom Geschlechtsstereotyp ist das Geschlechtsrollenstereotyp abzugrenzen. Dieses entstammt der sozialen Rolle, die beide Geschlechter typischerweise in der Gesellschaft einnehmen und beinhaltet sowohl deskriptive als auch präskriptive Aspekte. Es enthält Ansichten darüber, welche Charakteristiken Individuen besitzen müssen, wenn sie eine bestimmte soziale Rolle ausfüllen. Beispielsweise werden Frauen als liebevoll und freundlich charakterisiert, da sie typischerweise eher fürsorgende Rollen einnehmen. In dem Ausmaß, in dem diese Rollen einem sozialen Konsens unterliegen, werden sie Teil einer Kultur. Diese Tatsache führt in der Konsequenz zu einer Legitimation und Aufrechterhaltung der Rollenverteilungen (Diekmann, Eagly, Mladinic, & Ferreira, 2005; Eagly & Steffen, 1984).

Eine ganze Reihe von Studien zeigt, dass Männern eher Merkmale wie Dominanz, Aggression, Unabhängigkeit, Hartnäckigkeit, Ehrgeiz, Zielstrebigkeit und Wettbewerbsgeist zugeschrieben werden. Frauen

werden dagegen eher als warm, einfühlsam, hilfsbereit, sozial, sensibel, rücksichtsvoll, personenorientiert, passiv und wenig durchsetzungsfähig beschrieben (Sczesny, 2003a; Wänke et al., 2003). Diese Überzeugungen über kraftvolle und zielorientierte (*agentic*) Eigenschaften auf Seiten der Männer und gemeinschaftsorientierter (*communal*) Merkmale seitens der Frauen sind dabei ebenso deskriptiv wie präskriptiv zu verstehen (Eagly & Sczesny, im Druck).

Entgegen des hohen Konsens darüber, wie Frauen und Männer beschrieben werden, liefern Erhebungen ihrer tatsächlichen Eigenschaften ein weniger einheitliches Bild. Zwar zeigt sich tatsächlich, dass Frauen eine hohe Verträglichkeit, Wärme und Offenheit für Gefühle, Männer dagegen eher eine hohe Durchsetzungsfähigkeit und Offenheit für Ideen aufweisen, jedoch ist dabei die interindividuelle Varianz insgesamt höher als die Varianz zwischen den Geschlechtern (Costa Jr., Terracciano & McCrae, 2001).

Die den Frauen zugesprochenen Eigenschaften werden allgemein als sehr sympathisch angesehen, so dass Eagly & Mladinic (1994) sogar von einem „women-are-wonderful“-Effekt sprechen. Jedoch ist auch zu bemerken, dass eben diese Eigenschaften weniger mit beruflichem Erfolg, Leistungsfähigkeit und Kompetenz assoziiert sind (Wänke et al., 2003). Dagegen werden die Attribute, die in der Regel eher bei Männern wahrgenommen werden, gleichsam auch einem idealen Manager attestiert (Powell, Butterfield & Parent, 2002).

Dieses als „think manager - think male“ bekannt gewordene Phänomen (Schein, 1973, 1975) zeigt, dass über verschiedene Kulturen hinweg das Image einer erfolgreichen Führungskraft eher mit kraftvollen und maskulinen als mit femininen Eigenschaften verbunden wird, und zwar sowohl von weiblichen als auch von männlichen Befragten (Schein, 1973, 1975; Sczesny, 2003b; Sczesny, Spremann & Stahlberg, 2006). Einige Forschungsergebnisse deuten insofern einen Wandel an, als dass sich dieser Effekt bei weiblichen Befragten

abzuschwächen scheint, bei Männern dagegen konstant bleibt (Schein, Müller, Lituchy & Liu, 1996; Sczesny, 2003a).

Doch warum besteht dieses Stereotyp angesichts der veränderten gesellschaftlichen Rollen von Männern und Frauen weiterhin?

Eine mögliche Erklärung entstammt der „Subkategorien“- Theorie (Allport, 1954; Hewstone, 1994; Weber & Crocker, 1983). Diese geht davon aus, dass Frauen, die sich nicht entsprechend der weiblichen Rolle verhalten nicht das übergeordnete Stereotyp verändern, sondern neu gebildeten Unterkategorien zugeordnet werden. So fiel beispielsweise die Bundeskanzlerin von Deutschland Angela Merkel in eine Subkategorie „Karrierefrau“, die wiederum unabhängig vom regulären Frauenstereotyp durch eigene Eigenschaften und Verhaltensweisen charakterisiert ist (Wänke et al., 2003; Kunda & Oleson, 1995).

Eine alternative Erklärung sieht die Gründe in den sozialen Rollen von Männern und Frauen in der Gesellschaft (Eagly & Steffen, 1984). Heutige Geschlechtsrollenstereotype wurden gebildet, als die Rolle des Mannes als Ernährer mit Status, Macht und Durchsetzungskraft assoziiert wurde, die Frau dagegen vornehmlich in fürsorgenden und mütterlichen Rollen auftrat. Diese klassische Rollenverteilung hat sich in den letzten Jahren verändert, doch fand der Wandel größtenteils asymmetrisch statt (Diekmann & Eagly, 2000). Heute verdienen sehr viel mehr Frauen ihren eigenen Lebensunterhalt. Allerdings arbeiten sie dabei weitaus häufiger in Teilzeit und im Allgemeinen eher in von Frauen dominierten Berufsbereichen wie Verkauf, Sekretariat, Kranken- und Pflegeberufen. Hinzu kommt, dass Frauen immer noch mehr Pflichten im Haushalt wahrnehmen (Eagly & Sczesny, im Druck). Dagegen trat bei Männern keine Verlagerung des Tätigkeitsschwerpunktes von der Berufstätigkeit zu heimischen Pflichten ein (Diekmann et al., 2005).

Dieser einseitige Veränderungsprozess führt zu einer Wahrnehmung von Stabilität bezüglich der Rolle des Mannes bei gleichzeitiger Dy-

namisierung der Rolle der Frau (Diekmann & Eagly, 2000). So zeigt sich beispielsweise, dass maskuline Eigenschaften wie Entschlossenheit, Ehrgeiz, Intelligenz und Selbstsicherheit heute stärker zu den sozial erwünschten Eigenschaften einer Frau zählen als noch in den 80er Jahren (Diekmann et al., 2005; Swazina et al., 2004).

Die Dynamisierung traditioneller Rollen führt dazu, dass Frauen immer häufiger in Führungspositionen anzutreffen sind, dabei allerdings weltweit eher im mittleren Management als im Top-Management (Eagly, 2003; Eagly & Sczesny, im Druck; Hannover & Kessels, 2003). Die Tatsache, dass Frauen weitaus seltener Karriere machen als Männer liegt in einem Zusammenwirken sozialer und psychologischer Barrieren begründet.

Psychologische Ursachen liegen beispielsweise darin, dass Frauen oftmals selbst einem traditionell femininen Geschlechtsrollen-Selbstkonzept folgen und sich daher in Bewerbungssituationen vielfach unterschätzen, was wiederum ihre Karrierechancen negativ beeinflusst (Sieverding, 2003). Ein weiterer Grund ist, dass es bei der Beurteilung von Frauen, insbesondere im Hinblick auf Führungsqualitäten, zu Vorurteilen und Fehleinschätzungen kommt.

Laut Eagly und Karau (2002) gibt es zwei Formen von führungsbezogenen Vorurteilen gegenüber Frauen: Zum einen wird Frauen im Vergleich zu Männern weniger Führungspotenzial zugeschrieben. Zum anderen wird tatsächlich gezeigtes Führungsverhalten einer Frau negativer bewertet als jenes von einem Mann. Insbesondere männliche Personen bewerten Frauen in Bezug auf deren Führungsfähigkeiten im Vergleich zu Männern negativer als weibliche Personen (Sczesny, Bosak, Neff & Schyns, 2004; Eagly, Makhijani, Klonsky, 1992). Bei Frauen dagegen wandelt sich dieses Schema. Beispielsweise geben Frauen heute eher an, dass sie in gleichem Ausmaß wie Männer über die für eine Führungsposition notwendigen Charakteristiken verfügten (Jackson, Engstrom & Emmers-Sommer, 2007; Schein et al., 1996).

Auch wenn die klassische „think manager - think male“- Denkweise nachzulassen scheint und sich Vorbehalte gegen Frauen in Führungspositionen abschwächen, zählen sie nach wie vor zu den wichtigsten Karrierehemmern (Heilman, 2001; Schein, 2001; Wellington, Kropf & Gerkovich, 2003).

So hatte Timothy Leary wohl ein erhabenes Bild von Frauen und insbesondere von der Gesellschaft als er feststellte, dass ein Streben nach Gleichheit mit Männern bei Frauen lediglich einen Mangel an Ehrgeiz ausdrücke. Betrachtet man das Gesamtbild der aktuellen Forschungsergebnisse, so finden sich insgesamt wenige Hinweise auf starke Veränderungen der gesellschaftlichen Geschlechtsrollenstereotype (Eagly & Sczesny, im Druck). In einer Arbeitswelt, wie sie sich heute darstellt, könnten Frauen wohl kaum ein ambitionierteres Ziel anstreben als Gleichheit.

So lässt sich für die vorliegende Studie die Hypothese formulieren, dass sich Menschen bei der Einschätzung von Führungsstärke am biologischen Geschlecht der zu beurteilenden Personen orientieren und Männern eine höhere Führungsstärke zusprechen.

3.3. Das Attraktivitätsstereotyp oder „what is beautiful is good“

„Outside the realm of ideas, beauty rules.“

Nancy Etcoff (1999, S. 4)

Die Attraktivität eines Menschen liegt nicht, wie der Volksmund sagt, im Auge jedes Betrachters. Vielmehr sind sich Menschen, auch über verschiedene Kulturen und Altersgruppen hinweg, hinsichtlich der Attraktivität anderer so einig, wie sie es sonst selten sind. (Langlois, Kalakanis, Rubenstein, Larson, Hallam & Smoot, 2000).

Studien zufolge haben schon Neugeborene eine Präferenz für attraktive Gesichter und betrachten attraktive Gesichter länger als unattraktive Gesichter (Slater et al., 1998). Dies könnte als Hinweis auf ein

genetisch verankertes Schönheitsempfinden gewertet werden. Darüber hinaus zeigt die Forschung weiterhin, dass Menschen extrem schnell erkennen, ob jemand attraktiv ist. So waren Versuchspersonen in der Lage, bereits nach einer 100 Millisekunden dauernden Betrachtung eines Bildes anzugeben, wie attraktiv der abgebildete Mann oder die abgebildete Frau war (Locher, Unger, Sociedade & Wahl, 1993). Die Autoren schließen aus diesem Befund, dass das menschliche Wahrnehmungssystem automatisch und mühelos mit Attraktivität assoziierte Hinweise identifiziert und diese im Informationsverarbeitungsprozess Vorrang haben.

Studien zum deskriptiven Attraktivitätsstereotyp zeigen, dass schönen Menschen eine Vielzahl an positiven Eigenschaften und Charakterzügen attribuiert werden. Bei diesem als „what is beautiful is good“ (Dion, Berscheid & Walster, 1972) bekannt gewordenen Phänomen handelt es sich um einen Überlagerungseffekt, der in der Psychologie auch als Halo-Effekt bezeichnet wird. „Was schön ist, ist auch gut“ – Die Schönheit strahlt wie ein Heiligenschein (englisch: halo) über die Persönlichkeit eines Menschen. Die menschliche Wahrnehmung wird dahingehend verzerrt, dass schöne Menschen in fast allen Dimensionen günstiger beurteilt werden (Ebner, Gathmann & Wiedermann, 2002).

Das Schönheitsstereotyp und seine Konsequenzen sind heute gut dokumentiert. Die Ergebnisse vieler Studien zeigen, dass attraktive Menschen unter anderem als empfindsamer, freundlicher, interessanter, stärker, gelassener, bescheidener und offener wahrgenommen werden (Ebner et al., 2002). Weiterhin werden ihnen im Vergleich mit weniger attraktiven Menschen allgemein mehr Geselligkeit und soziale Kompetenz, Dominanz, ein höherer beruflicher Status sowie mehr beruflicher Erfolg, eine höhere Lebenszufriedenheit sowie eine glücklichere Beziehung zugeschrieben (Feingold, 1992). Langlois et al. (2000) zufolge schneiden schöne Menschen insbesondere in den Bereichen berufliche und akademische Kompetenz, beruflicher Er-

folg, mentale Einstellung, soziale Anziehungskraft und Popularität, interpersonale Kompetenz, Erfahrungen mit dem anderen Geschlecht und körperliche Gesundheit günstiger ab und werden auch dementsprechend vorteilhaft behandelt.

In einer Studie mit simulierten Bewerbungssituationen fanden Locher et al. (1993), dass attraktiven „Bewerbern“ sehr schnell eine bessere Passung mit der zu vergebenden Position und eine höhere Kooperationsbereitschaft attestiert wurden als weniger attraktiven Kandidaten.

Diese Forschungsergebnisse deuten ferner an, dass sich die genannten Zusammenhänge im Alltag im Sinne einer Sich-selbsterfüllenden-Prophezeiung bewahrheiten können (Snyder, Decker Tanke & Berscheid, 1977). So zeigt sich beispielsweise, dass attraktive Personen, insbesondere attraktive Männer, tatsächlich ein höheres Jahresgehalt verdienen als weniger attraktive Arbeitnehmer (Roszell, Kennedy & Grabb, 2001).

Das Konzept der Sich-selbsterfüllenden-Prophezeiung (Merton, 1948) geht davon aus, dass schöne Menschen in ihrer sozialen Umwelt bestimmten Erwartungen begegnen. Ihre Interaktionspartner verhalten sich entsprechend ihren Erwartungen in einer Art und Weise, die bei ihren attraktiven Gegenübern genau das vorab vermutete Verhalten provoziert. Um diese mögliche Erklärung zu testen, untersuchten Anderson, John, Keltner und Krings (2001) den Einfluss von Attraktivität auf den sozialen Status einer Person. Die Autoren prüften in rein weiblichen, rein männlichen und in gemischten Gruppen, welche Personen jeweils hohe Statuspositionen erlangten. Sie konnten zeigen, dass zumindest bei Männern eine positive Korrelation zwischen Attraktivität und sozialem Status bestand. Personen mit höherem Status wurden außerdem eher Führungsaufgaben zugetraut. Daher liegt die Vermutung nahe, dass attraktiven Menschen auch eine höhere Kompetenz im Bereich Führungsverhalten zugesprochen wird.

Tatsächlich konnten verschiedene Untersuchungen zeigen, dass das Aussehen einen Einfluss auf die Wahl eines politischen Führers hat

(Budesheim & De Paola, 1994; Efran & Patterson, 1974). Die jeweils attraktiveren Kandidaten wurden mit einer höheren Wahrscheinlichkeit gewählt als unattraktive Kandidaten. Auch bei der Auswahl von Führungspersonen in der Wirtschaft spielt die Attraktivität eine Rolle. Attraktiveren Bewerbern werden im Allgemeinen mehr Führungsqualitäten zugesprochen und ihnen werden eher Arbeitsstellen angeboten, in denen es auf Führungskompetenz ankommt (Sczesny et al., 2006; Sczesny & Kühnen, 2004).

Resultierend aus diesen Ergebnissen wird für die vorliegende Fragestellung die Hypothese aufgestellt, dass Attraktivität für die Wahrnehmung von Führungsstärke relevant ist und attraktiveren Personen eine höhere Führungsstärke zugeschrieben wird.

3.4. Das „Lack-of-fit“- Modell

Sowohl das Geschlechts- als auch das Attraktivitätsstereotyp haben für sich genommen in der Forschung bisher sehr viel Aufmerksamkeit erfahren. Aus diesen empirischen Ergebnissen heraus konnten die Hypothesen aufgestellt werden, dass bei Männern eine höhere Führungsstärke wahrgenommen werden sollte, genauso wie insgesamt bei attraktiveren Personen. Denkbar wäre allerdings auch, dass beide Stereotype nicht unabhängig voneinander wirken, sondern in einer bestimmten Kombination.

Das so genannte „Lack-of-fit“- Modell (Heilman, 1983) postuliert ein Zusammenwirken von Geschlecht und Attraktivität und gelangt somit - bezogen auf die hier behandelte Fragestellung - zu einer anderen Vorhersage als das „think manager – think male“- Phänomen und die „what is beautiful is good“- Heuristik.

Das „Lack-of-fit“- Modell erklärt die Effekte von Geschlecht und Attraktivität auf wahrgenommene Fähigkeiten. Kernaussage dieses Konzepts ist, dass im organisationalen Kontext bei attraktiven Frauen die wahrgenommene Passung (*fit*) zwischen Person und Füh-

rungsposition besonders gering ausgeprägt ist. Dies wird folgendermaßen begründet:

Ob ein Bewerber für eine offene Stelle ein Vertragsangebot erhält oder nicht hängt sehr stark davon ab, ob die entscheidende Person eine Kongruenz zwischen den Fähigkeiten und Eigenschaften des Bewerbers und den Anforderungen der beruflichen Position erlebt. Je größer die Diskrepanz zwischen diesen beiden Wahrnehmungen, desto eher wird ein Scheitern des Bewerbers in der Position angenommen und damit fällt die Auswahlentscheidung wahrscheinlich negativ aus.

Nun zeigt die Forschung zum „think manager – think male“- Phänomen, dass Führungspositionen eher mit maskulinen, kraftvollen Attributen assoziiert werden (Powell et al., 2002; Schein, 1973; Sczesny, 2003b; Sczesny et al., 2006).

Bereits aus diesem Grund sollten in vielen Fällen die Eigenschaften, die eine Bewerberin mitbringt, als nicht förderlich für den Erfolg in einer Führungsposition angesehen werden. Hinzu kommt, dass Attraktivität bei Frauen in der Regel mit Femininität einhergeht, bei Männern dagegen mit Maskulinität (Gillen, 1981).

Bei Bewerbungen um eine Führungsposition konnte - zumindest im Labor - gezeigt werden, dass attraktive Männer und durchschnittlich attraktive Frauen tatsächlich mit einer höheren Wahrscheinlichkeit ein Angebot bekamen als sehr attraktive Frauen und weniger attraktive Männer. Attraktivität ist somit nur für Männer ein Vorteil, für Frauen dagegen eher ein Nachteil. Dieses Muster konnte jedoch nur im Hinblick auf Führungspositionen gefunden werden. Galt die Bewerbung einer Bürotätigkeit, eine Arbeit, die weder eindeutig maskulin noch eindeutig feminin wahrgenommen wird, war Attraktivität für beide Geschlechter von Vorteil (Heilman & Saruwatari, 1979).

Heilman und Stopeck (1985) konnten darüber hinaus zeigen, dass in Führungspositionen das gleiche Verhalten von attraktiven verglichen mit unattraktiven Frauen negativer bewertet wird und sich dies auch auf Beförderungen und Gehaltserhöhungen auswirkt. In dieser Stu-

die erhielten Versuchspersonen Beurteilungen von Vorgesetzten über verschiedene Arbeitnehmer. Die Gesamtbeurteilung war dabei in jedem Fall gleich gut. Lediglich die berufliche Position und das Aussehen der bewerteten Arbeitnehmer unterschieden sich. Die Arbeitnehmer wurden entweder als Auszubildende/r als Bürokaufmann/frau vorgestellt oder als Teilnehmerin/in eines Führungskräftenachwuchsprogramms. Weiterhin fand sich auf dem ausgehändigten Bogen ein Foto einer (un)attraktiven Frau oder eines (un)attraktiven Mannes.

Die Aufgabe der Versuchspersonen bestand erstens darin, die Arbeitnehmer auf verschiedenen Dimensionen zu bewerten, darunter deren Führungsfähigkeiten (z.B. Innovationsverhalten, Organisationsfähigkeit, Motivationsfähigkeit) und persönliche Erfolgswirksamkeit (z.B. emotionale Stabilität, positive Einstellung, Kooperationsfähigkeit). Zweitens sollten sie bei jedem Arbeitnehmer die Angemessenheit einer Gehaltserhöhung sowie einer Beförderung angeben. Die Ergebnisse zeigen, dass attraktive Frauen insgesamt besser bewertet wurden, wenn sie bei gleicher Leistung eine Ausbildung zur Bürokauffrau absolvierten als wenn sie als Führungsnachwuchskraft vorgestellt wurden. Gehaltserhöhungen und auch Beförderungen schienen den Beurteilern angemessener bei attraktiven Bürokauffrauen und bei unattraktiven weiblichen Führungsnachwuchskräften.

Während diese Studien ein homogenes Bild des Zusammenwirkens von Geschlecht und Attraktivität für Frauen liefern, sind die Ergebnisse bezogen auf Männer weniger eindeutig. Gillen (1981) und Heilman und Sarutawari (1979) fanden einen positiven Zusammenhang mit Attraktivität, in der Studie von Heilman und Stopeck (1985) hatte Attraktivität bei Männern jedoch weder einen Einfluss auf die Gesamtbeurteilung des Führungsverhaltens noch auf die empfohlenen Entwicklungsmaßnahmen wie Gehaltserhöhung oder Beförderung.

Resultierend aus dem „Lack-of-fit“-Modell wird für die vorliegenden Studien die Hypothese aufgestellt, dass eine Interaktion zwischen der

Attraktivität und dem Geschlecht der zu beurteilenden Personen auftritt, so dass attraktiven Frauen im Vergleich zu weniger attraktiven Frauen eine niedrigere Führungsstärke, attraktiven Männern im Vergleich zu weniger attraktiven Männern dagegen eine höhere Führungsstärke zugeschrieben wird.

3.5. Schlussfolgerungen für die Wahrnehmung von Führungsstärke

Zusammenfassend bleibt festzuhalten, dass Menschen bei der Einschätzung anderer Menschen Heuristiken und Stereotype anwenden, die eine schnelle und ökonomische Orientierung bieten, jedoch auch zu systematischen Urteilsverzerrungen führen können. Wichtig in diesem Kontext ist, dass das Hauptaugenmerk der sozialen Kognitionsforschung bis heute auf dem *Prozess* der Urteilsbildung liegt (Krueger & Funder, 2004). Die dokumentierten Fehler und Verzerrungen stellen allesamt Abweichungen von einem normativ als korrekt angenommenen Urteilsprozess dar. Nicht weil Denk- und Entscheidungsprozesse zu schlechten sozialen Entscheidungen führen gelten sie als falsch, sondern weil sie unter Laborbedingungen von logischen Inferenzmodellen abweichen (Funder, 1987). Die kognitive Sozialpsychologie liefert immer wieder frappierende Ergebnisse. Doch ihre Methodologie ist von ökologischer Validität meist weit entfernt.

Experimentell werden starke Kontrastierungen und Simplifizierungen („Strichmännchen“, Gesichts- oder Körperschemata, von Computern generierte Prototypen von Gesichtern, Durchschnittsgesichter, usw.), künstliche Stimuli oder hypothetische Personen genutzt, um durch minimale Veränderungen des Stimulusmaterials, des Versuchablaufes oder der experimentellen Bedingung jeweils unterschiedliche Prozesse und Effekte zu induzieren. Auf diese Art und Weise kann genau untersucht werden, welche Veränderungen und welche Eigenschaften welche Auswirkungen beispielsweise im Hinblick auf die Bildung eines Eindrucks haben.

Diese Effekte haben einen hohen Aufklärungswert, stellen jedoch bestenfalls eine Annäherung an wirklichkeitsnahe Prozesse dar. Beispielsweise konzentriert sich der Großteil der existierenden Stereotypforschung auf Fehlwahrnehmungen und Verzerrungen, die aus der Anwendung von Stereotypen resultieren, wobei eine genaue Analyse der Verarbeitung sozialer Informationen im Vordergrund steht (Judd & Park, 1993). So liegt das Hauptaugenmerk der Forschung zu Geschlechts- und Attraktivitätsstereotypen auf Fragen wie: Welchen Inhalts sind diese Stereotype? Wie und wann verändern sich Stereotype und stereotype Rollenzuschreibungen? Unter welchen Umständen und aus welchen Gründen verändern sie sich gerade nicht? Wie kann die Aktivierung von Stereotypen reduziert werden?

Jede dieser Fragen ist wichtig und interessant. Ihre Antworten und die Schlussfolgerungen, die sie erlauben, sind von essentieller Bedeutung für das soziale Miteinander.

Die Genauigkeit von Personenwahrnehmung jedoch ist und war nie erklärtes Ziel der kognitiven Sozialpsychologie (Funder, 1987), genauso wenig wie die Frage, ob Stereotype im Großen und Ganzen zutreffen oder nicht zutreffen (Judd & Park, 1993). Grundsätzlich wird die Genauigkeit von sozialen Urteilen von den klassischen Vertretern des „*heuristics and biases*“- Ansatzes nicht ausgeschlossen (Kahneman, Slovic & Tversky, 1982; Kahneman & Tversky, 1996). Ob sie damit jedoch richtig oder falsch liegen, kann die Forschung innerhalb dieses Ansatzes nicht beantworten, da diese Frage typischerweise nicht gestellt wird.

Insofern lässt sich festhalten, dass die Frage, ob die Anwendung von Heuristiken und Stereotypen zu guten sozialen Einschätzungen führt, bisher weitgehend unerforscht blieb. Daher kann in Bezug auf die Frage nach dem Ergebnis der Personenwahrnehmung, das heißt ob Menschen die Führungskraft anderer zutreffend einschätzen können, aus der sozialen Kognitionsforschung heraus keine gerichtete Hypothese aufgestellt werden.

Anstelle der Frage, ob ein Urteil richtig oder falsch ist, steht in dieser Forschungslinie die Frage im Vordergrund, auf welche Art und Weise ein Urteil zustande kommt (Funder, 1999). Da Eindrucksbildungsprozesse in diesem Rahmen extensiv erforscht wurden, können für den Prozess der Wahrnehmung von Führungsstärke eindeutige, jedoch konkurrierende Hypothesen aufgestellt werden.

Gemäß der Forschung zum „think manager – think male“ (Powell et al., 2002; Schein, 1973; Sczesny, 2003b; Sczesny et al., 2006) sowie zum „what is beautiful is good“ Stereotyp (Dion et al., 1972; Feingold, 1992; Langlois et al., 2000) ist anzunehmen, dass Geschlecht und Attraktivität jeweils einen direkten Einfluss auf die Wahrnehmung von Führungsstärke haben sollten, so dass Männern und attraktiveren Personen eine höhere Führungsstärke zugeschrieben würde.

Entgegen dieses direkten Einflusses würde das „Lack-of-fit“-Modell (Heilman, 1983; Heilman & Stopeck, 1985) eine Interaktion beider Merkmale vorhersagen, so dass attraktiven Männern eine hohe, attraktiven Frauen jedoch eine geringere Führungsstärke zugeschrieben würde.

Stellt man über den Prozess der Eindrucksbildung hinaus dennoch Vermutungen über die Validität intuitiver Persönlichkeitseinschätzungen an, so lautet die wohl wahrscheinlichste Annahme, dass die Wahrnehmung der intuitiven Beurteiler durch die Anwendung von Heuristiken und Stereotypen verzerrt ist (Funder, 1987). Dies wäre in der vorliegenden Studie der Fall, wenn nur das Geschlecht und die Attraktivität der zu Beurteilenden als Anhaltspunkte für deren Führungsqualitäten genommen würden und die tatsächliche Führungsstärke nicht erkannt würde. Grundsätzlich denkbar wäre allerdings auch, dass die Beurteiler zwar sowohl das Geschlechts- als auch das Schönheitsstereotyp verwenden, darüber hinaus jedoch zusätzlich die tatsächliche Führungsstärke erkennen. Diese Alternative würde bedeuten, dass der Eindruck zwar mit einer Irrtumskomponente behaftet wäre, obendrein jedoch auch Genauigkeit beinhaltet.

Da jedoch neben der Frage nach dem Beurteilungsprozess auch die Frage nach der Validität der Einschätzungen von Führungsstärke untersucht werden soll, wird im Folgenden ein Forschungsansatz vorgestellt, der diesbezüglich Schlussfolgerungen zulässt.

4. Personenwahrnehmung aus evolutionspsychologischer Perspektive

Menschen sind hochgradig soziale Wesen. Für die menschliche Evolutionsgeschichte bedeutet dies, dass viele Probleme im Kontext von Überleben und Fortpflanzung sozialer Natur waren (Haselton & Funder, 2006). Aus diesem Grunde sind viele Aspekte des menschlichen Denkapparates angepasst an die wiederkehrenden Probleme, genauso wie an die Möglichkeiten des sozialen Mit- und Gegeneinanders. Das evolutionspsychologische Augenmerk liegt dabei besonders auf der Tatsache, dass Denken, Fühlen und Entscheiden einzig und allein einer Funktion dient: Handeln (Kenrick, Schaller & Simpson, 2006). Die spezifischen Gedanken und Gefühle, die in einem gegebenen sozialen Kontext erlebt werden, sind sehr wahrscheinlich eben jene, die in der evolutionären Vergangenheit zu spezifischen Verhaltensweisen führten, die im Regelfall der reproduktiven Fitness zuträglich waren, die also dem aus evolutionärer Perspektive übergeordneten Ziel, nämlich der Weitergabe der eigenen Gene, dienlich waren (Schaller, Park & Kenrick, 2007).

Als Produkt dieses Selektionsdruckes sollten sich kognitive Prozesse durchgesetzt haben, die schnelle und adaptive Verhaltensantworten ermöglichen, sobald fitnessrelevante Informationen in der Umgebung entdeckt werden. In einem hochgradig sozialen Leben sind es meist andere Personen, die fitnessrelevante, das heißt im Kontext von Überleben und Reproduktion wichtige Herausforderungen darstellen. Daher sollten Menschen diesen mit besonderer Aufmerksamkeit begegnen.

Soziale Herausforderungen lassen sich in zwei wesentliche Kategorien einteilen. A) Soziale Aussichten und Möglichkeiten, deren Erlangung sich positiv auf die Fitness auswirkt und B) Gefahren, deren erfolgreiche Vermeidung sich positiv auf die Fitness auswirkt (Schaller et al., 2007).

Bezogen auf die Vermeidung von Gefahren sollte ein beträchtlicher Selektionsdruck nicht nur auf die Genauigkeit, sondern auch auf die Schnelligkeit der Wahrnehmung und Eindrucksbildung gewirkt haben. Jemand, der gefährliche Menschen, Tiere und Objekte in der Umwelt nicht sofort als solche identifizieren konnte, hatte beträchtliche Folgeschäden zu tragen. Bezüglich des Erkennens von Risiken aus der Umwelt gibt es eine ganze Reihe von Forschungsarbeiten, die demonstrieren, dass Menschen bedrohliche Stimuli wie Schlangen und Schusswaffen schneller erkennen als neutrale Objekte wie Blumen oder Toaster (Fox, Griggs & Mouchlianitis, 2007; Öhman & Mineka, 2001). Weiterhin zeigt sich, dass Menschen Wut und Ärger bei anderen Menschen zuverlässig einschätzen können und zornige von freundlichen Gesichtern schnell differenzieren können. Dies ist selbst dann noch möglich, wenn nur die Augenregion einer Person zu sehen ist (Fox & Damjanovic, 2006). Das visuelle System des Menschen scheint darauf ausgerichtet zu sein, Bedrohungen schon auf Basis von extrem wenigen Informationen zu erkennen. So zeigte eine Studie von Bar, Neta und Linz (2006), dass Gesichter bereits nach 39 Millisekunden übereinstimmend als bedrohlich oder harmlos eingestuft wurden.

Wahrnehmungsgenauigkeit und -schnelligkeit sind jedoch nicht nur bei der Vermeidung von sozialen Gefährdungen wichtig. Eine mögliche positive Aussicht von sozialen Interaktionen ist die Wahl eines Partners. In diesem Kontext gilt es zunächst die Kandidaten zu identifizieren, die erstrebenswerte Eigenschaften aufweisen, um dann eine schnelle Entscheidung darüber zu treffen, ob es sich lohnt in diese Möglichkeit zu investieren oder ob die Suche nach anderen potenziel-

len Partnern die bessere Alternative darstellt. In diesem Kontext zeigt sich, dass Menschen ein spezielles Geschick darin haben, gute und anziehende potenzielle Partner rasch zu identifizieren (Locher et al., 1993; Maner et al., 2003).

Viele Gegebenheiten des sozialen Lebens sind allerdings nicht pauschal der einen oder anderen Kategorie sozialer Herausforderungen zuzuordnen. Vielmehr beinhalten sie häufig beides, Gefahren und gleichzeitig Möglichkeiten. Ein solches Beispiel stellt die Einschätzung der Führungsstärke fremder Personen dar.

4.1. Erkennen von Führungsstärke aus evolutionspsychologischer Perspektive

Das Merkmal Führungsstärke ist in starkem Maße mit der Erlangung und Aufrechterhaltung von Status in sozialen Gruppen assoziiert, unabhängig davon, ob es sich bei der Gruppe um eine Jagdkoalition, Freizeitclique oder eine neue Projektgruppe bei der Arbeit handelt.

Soziale Integration und insbesondere der relative soziale Status innerhalb einer Gruppe haben immense Auswirkungen auf die individuelle Überlebens- und Fortpflanzungswahrscheinlichkeit. Neben Vorteilen wie Schutz, Ressourcen, Arbeitsteilung, Spezialwissen, usw., die Gruppen gewöhnlich bieten, verschaffen sie vor allem Zugang zu potenziellen Partnern. Grundsätzlich haben von der Gruppe ausgeschlossene Individuen geringe Möglichkeiten, überhaupt einen Partner zu finden. Personen mit einem niedrigen Status treffen allein aufgrund ihrer Zugehörigkeit zu einer Gruppe eher potenzielle Partner, verfügen jedoch in der Regel nicht über genügend Ressourcen, um erfolgreich um diese werben oder sie dauerhaft binden zu können (Schaller, 2008).

Statusverhandlungen und der damit verbundene Zugang zu Ressourcen sind daher Kernaspekte der Bildung und Veränderung sozialer Gruppen (Buss, 2004). Da der relative Status in der Gruppenhierar-

chie von so entscheidender Bedeutung ist, sollten Menschen sensibel gegenüber Hinweisen auf den eigenen Rang in einer Gruppe sein.

So zeigen verschiedene Forschungsergebnisse nicht nur, dass sich bei der Neubildung von Gruppen extrem schnell, teilweise innerhalb nur einer Minute, Gruppenhierarchien bilden (Fisek & Ofshe, 1970), sondern auch, dass die neuen Gruppenmitglieder ihren eigenen zukünftigen Status in der Gruppe genau einschätzen konnten, sobald sie die anderen Gruppenmitglieder gesehen hatten und noch bevor ein Wort gesprochen wurde (Kalma, 1991).

Anderson, Srivastava, Beer, Spataro und Chatman (2006) führten beispielsweise eine Reihe von Studien durch mit deren Hilfe sie nachweisen konnten, dass Menschen in realen Gruppen ihren Status sehr genau einschätzen können. Dieses Ergebnis trat unabhängig von Geschlecht und von der zu bearbeitenden Aufgabe auf und auch, wenn die Gruppe lediglich 20 Minuten miteinander interagiert hatte und die einzelnen Teilnehmer somit im Wesentlichen noch Fremde waren. Weiterhin bezogen sich ihre korrekten Einschätzungen sowohl auf objektive Anzeichen von Status, beispielsweise wie hoch ihr Redeanteil war, als auch auf weniger eindeutige Kennzeichen, zum Beispiel wie sehr sie von den anderen Gruppenmitgliedern respektiert wurden.

Über die adäquate Einschätzung des eigenen Ranges in der Gruppe hinaus, sollten Menschen anderen Gruppenmitgliedern gegenüber besonders aufmerksam sein. In Gruppen können nur wenige Personen viel Einfluss haben und daher steigt unweigerlich jemand in der Hierarchie ab, wenn jemand anderes aufsteigt (Anderson et al., 2006). Wirkt sich also das Verhalten einer anderen Person auf deren relativen Status aus, hat dies in der Folge ebenfalls einen Effekt auf die eigene Rangstufe. Die Aktivitäten anderer Personen sollten daher aufmerksam verfolgt werden. Dies gilt sowohl für bestehende Gruppenmitglieder, genauso aber für unbekannte Menschen, potenzielle neue Gruppenmitglieder. Diese können sich entweder als irrelevant,

hilfreich oder bedrohlich im Hinblick auf den eigenen sozialen Status erweisen.

Besitzt das neue Mitglied Führungsstärke, so hat es das Potenzial, einen hohen oder sogar höheren sozialen Status in der eigenen Gruppe zu erreichen. Diese Person bietet daher entweder die Möglichkeit für eine Allianz zur Erreichung eines höheren Status oder sie stellt eine Bedrohung für den eigenen Status dar. Alternativ dazu könnte der Fremde kein Potenzial aufweisen, Rang und Macht in der eigenen Gruppe zu erlangen und deswegen in dieser Hinsicht von keiner weiteren Bedeutung sein. In welche Kategorie der einzelne Gegenüber auch fallen mag, die valide Einschätzung seiner Führungsstärke und daran angekoppelt das eigene Verhalten als Reaktion auf diese Einordnung, entscheidet über den zukünftigen Rang innerhalb der sozialen Gruppe.

Gegenüber dieser essentiellen Bedeutung von Wahrnehmungsgenauigkeit ist es als kaum wahrscheinlich anzusehen, dass die soziale Evolution einen Denkkapparat geschaffen hat, der grundlegend anfällig ist für Fehler und Irrtümer. Das massive und komplexe Informationsverarbeitungssystem des Menschen sollte vielmehr speziell darauf ausgelegt sein, in sozialen Situationen und Umgebungen gute Einschätzungen und Entscheidungen zu treffen (Schaller et al., 2007). Wichtig in diesem Zusammenhang ist die Feststellung, dass kein Evolutionspsychologe behaupten würde, Menschen lägen mit ihren Einschätzungen und Urteilen immer genau richtig. Vielmehr sollten Menschen in der Lage sein, ihre Mitmenschen schnell relativ akkurat - wenn auch nicht perfekt - in einzuschätzen.

Da Persönlichkeitseigenschaften überdauernde Verhaltensdispositionen darstellen, sollte eine bessere Einschätzung der Persönlichkeit zu besseren Verhaltensvorhersagen führen und damit bessere soziale Entscheidungen erlauben (Haselton & Funder, 2006). In diesem Gedanken wird erneut der Grundgedanke des evolutionären Ansatzes deutlich: Soziale Wahrnehmung dient sozialem Handeln. Sie steuert

adaptives Verhalten und liefert die Grundlage für angemessene soziale Reaktionen. Nur so sind Menschen in der Lage zu erkennen, wann Vermeidung, wann Annäherung angebracht ist, wen sie um eine Verabredung bitten, als Freund wählen oder mit wem sie sich besser nicht anlegen sollten.

Die Bedeutung guter Persönlichkeitseinschätzungen ist für das soziale Miteinander und eigene Fortkommen im Leben so fundamental, dass Haselton und Funder (2006) die Existenz eines Persönlichkeitseinschätzungs-Instinkts postulieren. Unter den Begriff Instinkt fällt hier eine Reihe von evolvierten, relativ autonomen und hochgradig spezialisierten Modulen, deren Zusammenwirken auch nach kurzen Interaktionen Inferenzen erlauben, die weit über die gegebenen Informationen hinausgehen. Persönlichkeitseinschätzungen scheinen dabei durch mindestens drei Merkmale charakterisiert zu sein. Zu diesen gehört Schnelligkeit, Unausweichlichkeit und (bis zu einem gewissen Grad) Unzugänglichkeit für das Bewusstsein. Als Adaptation an eine intensiv soziale Lebensweise sollte dieser Instinkt insbesondere dann schnelle und zutreffende Eindrücke liefern, wenn es um die Lösung basaler adaptiver Probleme geht. Dazu gehört die akkurate Einschätzung des Partnerwertes und der Verfügbarkeit eines potenziellen Partners, aber auch die Einschätzung des sozialen Status anderer.

Aus der Hypothese des Persönlichkeitseinschätzungs-Instinktes lassen sich verschiedene Hypothesen ableiten, von denen einige bereits eine empirische Basis haben, andere bisher weitgehend spekulativ bleiben (für einen Überblick siehe Haselton & Funder, 2006). Beispielsweise lautet eine These, dass diese Fähigkeit ohne explizites Training auftreten sollte. Empirisch zeigt sich in der Tat, dass bereits im frühen Kindesalter ein Verständnis von Persönlichkeit, spontane Persönlichkeitseinschätzungen und ein Verständnis davon, dass verschiedene Menschen aufgrund ihrer Eigenschaften auf das gleiche Ereignis unterschiedlich reagieren, auftreten.

Eine weitere Hypothese lautet, dass alle Menschen von Natur aus kompetent darin sein sollten, rasche und generell valide Einschätzungen anderer Menschen vorzunehmen, auch wenn sie diese kaum oder gar nicht kennen. In diesem Kontext sichtete Funder (1999) die bisherigen Erkenntnisse auf der Suche nach dem „good judge“, das heißt nach Eigenschaften, die einen Menschen mit besonders guter Menschenkenntnis ausstatten, und kam zu dem Schluss, dass diese bisher weitgehend erfolglos verlaufen sei.

Auch aktuellere Ergebnisse liefern ein inkonsistentes Bild. So fanden Vogt und Colvin (2003), dass Individuen mit einer hohen Gemeinschafts- und interpersonellen Orientierung bessere Einschätzungen gelangen. Weiterhin fanden sie einen Geschlechterunterschied in der Güte der Einschätzungen. Diese Ergebnisse konnten von Christiansen, Wolcott-Burnam, Janovics, Burns und Quirk (2005) nicht bestätigt werden. Sie fanden in ihrer Studie keinen Vorteil der Beurteiler mit interpersonell orientierten Eigenschaften wie Extraversion, Verträglichkeit und Gewissenhaftigkeit. Ihre Ergebnisse verwiesen dagegen auf einen positiven Effekt von Offenheit und allgemeinem Intelligenzniveau. Dieser Befund wiederum entspricht den Ergebnissen von Lippa und Dietz (2000), die einen positiven Zusammenhang zwischen der Einschätzungsgüte und allgemeinem Intelligenzniveau fanden, darüber hinaus jedoch weder einen Geschlechterunterschied noch eine Relation mit anderen Persönlichkeitsvariablen. Insofern lässt sich festhalten, dass bis heute die Festlegung auf bestimmte für die Menschenkenntnis förderliche Eigenschaften nicht gelang und somit allen Menschen schnelle und zuverlässige Einschätzungen gelingen sollten.

Schnelle Eindrücke von anderen Menschen müssen dabei nicht vollkommen korrekt sein. Vielmehr sind sie auch dann adaptiv, wenn sie fehlbar und unvollkommen sind, so lange sie zumindest minimal diagnostisch sind, so dass bereits auf dieser Basis adaptive Verhaltenskonsequenzen folgen können, statt erst nach längerem Überden-

ken und Zögern (Schaller, 2008). Von einer Möglichkeit zur empirischen Überprüfung der Diagnostizität erster Eindrücke handelt der folgende Abschnitt.

4.2. Empirische Überprüfung: das „Thin Slices“ – Paradigma

Die Genauigkeit von sozialer und insbesondere interpersoneller Wahrnehmung wurde in der empirischen Forschung seit den 20er Jahren intensiv behandelt. Eine Vielzahl unterschiedlicher Studien untersuchte die Präzision von subjektiven Urteilen, indem sie die Einschätzungen von Probanden mit externalen Kriterien verglich. Der häufigste Ansatz bestand darin, Mitglieder einer Gruppe sich gegenseitig einschätzen zu lassen und anschließend diese Angaben auf Übereinstimmung mit der Selbstauskunft zu prüfen (z.B. Adams, 1927; Dymond, 1950; für einen Überblick siehe Taft, 1955).

Dieses lebhaftes Interesse an der Frage nach Genauigkeit verschwand durch massive Kritik an den damaligen Methoden und Ergebnissen durch Cronbach (1955) nahezu vollkommen. Er demonstrierte, dass die gefundenen Übereinstimmungen zu einem ungewissen Grad durch verschiedene Einflüsse (z.B. stereotype Genauigkeit, geteilte Antworttendenzen, vermutete Ähnlichkeit, etc.) verursacht wurden und somit nicht unbedingt Aussagen über wahre Einschätzungsgenauigkeit trafen. Seine Intention bestand darin, einen besseren Umgang mit derart komplexen Daten anzuregen. Stattdessen galt die Erforschung der Wahrnehmungsgenauigkeit plötzlich für viele Wissenschaftler als verrufen (Funder, 1999) und der empirische Fokus verlagerte sich auf die Analyse von Prozessen.

Durch eine intensive Auseinandersetzung mit Cronbachs Kritik in den 80er und 90er Jahren und die dadurch angeregten methodischen Modifikationen (Funder, 1980; Harackiewicz & DePaulo, 1982; Kenny 1991), erwachte die Frage nach der sozialen Wahrnehmungsfähigkeit erneut. Dies geschah maßgeblich durch ein Forschungsparadigma,

dessen expliziter Schwerpunkt auf der Untersuchung der Validität des menschlichen Urteils liegt. Dieses untersucht mit Hilfe von sehr kurzen Beobachtungszeiträumen (*thin slices*) insbesondere die Genauigkeit von intuitiven Eindrücken über gänzlich fremde Menschen (*zero acquaintance situations*).

„Thin slices“ sind kurze Ausschnitte dynamischen Verhaltens. Die Verwendung von Transkripten als Stimulusmaterial ist eine Möglichkeit, meist werden jedoch kurze Video- oder Audiosequenzen eingesetzt. Die Dauer der Verhaltenssequenzen liegt gewöhnlich zwischen 3 Sekunden und 4 Minuten, jedoch nie über 5 Minuten. Sie enthalten Informationen eines Kommunikationskanals oder kombinieren mehrere, beispielsweise kann entweder nur das Gesicht oder aber Gesicht, Körper und Stimme gleichzeitig gezeigt werden. Ein methodischer Vorteil der „Thin slices“ besteht in der Möglichkeit, gezielt bestimmte Informationen reduzieren beziehungsweise vermeiden zu können. Beispielsweise kann der Sprachinhalt, die Vergangenheit der zu beurteilenden Person oder die Situation, in der das Verhalten stattfand ausgeblendet werden. So kann systematisch der Einfluss der einzelnen Kommunikationskanäle oder deren Wechselwirkung untersucht werden (Ambady, Bernieri & Richeson, 2000).

In einem prototypischen „Thin slices“- Experiment sehen ungeschulte Beobachter (im Folgenden als intuitive *Beurteiler* bezeichnet) kurze Videosequenzen der zu beurteilenden Personen (im Folgenden als *Zielpersonen* bezeichnet) und beurteilen diese jeweils direkt im Anschluss im Hinblick auf bestimmte Eigenschaften. Diese Sequenzen können unterschiedliches Verhalten zeigen. Eine typische Situation ist das Vorstellen in die Kamera. Aber auch andere Situationen wurden bereits verwendet, wie das Ablesen von Texten (Borkenau & Liebler, 1993), Mimen eines TV-Moderators (Lippa & Dietz, 2000), Kennen lernen eines Kommilitonen oder ein Gespräch mit einem guten Freund (Vogt & Colvin, 2003).

Eine große Anzahl von Forschungsarbeiten, die in den letzten beiden Jahrzehnten unter Rückgriff auf die „Thin slices“- Methode durchgeführt wurden, konnte die Validität der Personenwahrnehmung auf Basis von teilweise sehr kurzen Verhaltensbeobachtungen konsistent untermauern. Den Beurteilern gelangten unter anderem überzufällig genaue Einschätzungen über die Zielpersonen im Hinblick auf kurzfristige Gemütszustände wie zum Beispiel Emotionen, aber auch chronische und langfristige wie Depression oder Ängstlichkeit (Waxer, 1976, 1977), Persönlichkeitsdimensionen wie Extraversion und Geselligkeit (Kenny, Albright, Malloy & Kashy, 1994), die Tendenz zu altruistischem Verhalten (Fetchenhauer, Groothuis & Pradel, under review), Intelligenz (Murphy, Hall & Colvin, 2003), interpersonelle Rollen, interpersonelle Ziele wie beispielsweise das Ziel, einen bestimmten Eindruck zu hinterlassen, soziale Beziehungen wie Status- und Machthierarchien, Dominanz, Verwandtschaft, Bekanntschaft, Liebesbeziehungen oder Art und Qualität einer Beziehung. Auch in Bezug auf individuelle Leistungen konnten gute Voraussagen getroffen werden, wie die Evaluation von Lehrern durch ihre Schüler am Ende eines Schuljahres, ärztliche Effektivität bei der Heilung von Patienten, die Wahrscheinlichkeit, mit der ein Chirurg nach einer Behandlung verklagt wird und vieles mehr (für einen Überblick siehe Ambady et al., 2000).

Die Ergebnisse dieses Forschungsprogramms unterstützen somit die Annahme, dass Menschen grundsätzlich in der Lage sind, relativ akkurate Urteile über andere Menschen zu fällen. In einer Meta-Analyse berechneten Nalini Ambady und ihre Kollegen (2000) eine durchschnittliche Korrelation von $r = .27$ zwischen den Intuitivbeurteilungen und dem jeweiligen Referenzwert. Dieser Zusammenhang macht deutlich, dass es den Beurteilern zwar nicht gelang, die gefragte Persönlichkeitseigenschaft oder das Verhalten exakt vorherzusagen, jedoch besser als rein zufällig. Dies gilt auch dann, wenn Beurteiler und Zielperson zuvor kaum oder gar nicht miteinander interagiert haben und auch wenn die Dauer der Beobachtung extrem

kurz ist. Doch wie erklären Vertreter des „Thin slices“- Ansatzes diese Ergebnisse? Worauf basiert die Genauigkeit der intuitiven Beurteiler?

4.2.1.Kernannahmen des „Thin slices“- Paradigmas

Eine grundlegende Annahme dieses Paradigmas besteht darin, dass durch Verhalten nonverbale Informationen transportiert werden. Das nonverbale Verhalten jedes Individuums ist maßgeblich von seiner Persönlichkeit, Kultur und Lebenssituation geprägt und enthält somit wichtige Hinweise auf Affekte und Emotionen, Persönlichkeit, Überzeugungen, intrinsische Ziele und Motive und auch über die sozialen Beziehungen eben jener Person (Ambady et al., 2000).

Nonverbales Verhalten ist im Gegensatz zu verbalem Verhalten sehr schlecht zu unterdrücken beziehungsweise zu kontrollieren. Vor allem ist es von außen sichtbar, das heißt leicht zu beobachten, und damit besonders zugänglich für andere Menschen (DePaulo, Blank, Swaim & Hairfield, 1992). Demzufolge sollte die Beobachtung von Körpersprache, Mimik, Gestik, etc. den Beobachter mit besonders ertragreichen Informationen über den internalen Zustand des Beobachteten versorgen und ihn gleichzeitig zu validen Urteilen über die handelnde Person befähigen.

Als Illustration soll eine Studienreihe von Babad, Bernieri und Rosenthal (1989) dienen, die das Auftreten von Lehrern in Abhängigkeit von ihrer persönlichen stereotypen Verzerrungstendenz untersuchten. Zunächst teilten sie die Lehrer in zwei Gruppen mit hoher beziehungsweise niedriger Neigung zu ethnischen und sozioökonomischen Stereotypen ein. In vorherigen Untersuchungen wurden die Lehrer aus der Gruppe mit hoch ausgeprägter stereotyper Tendenz von ihren Schülern und Studenten eher als rigide, autokratisch, instabil, impulsiv und aufgabenorientiert, dafür als weniger personenorientiert und vertrauensvoll beschrieben als Lehrer aus der weniger vorurteilsbehafteten Gruppe. Die kombinierte Analyse von Transkripten plus 10sekündiger Videosequenzen, in denen die Lehrer entweder über ihre Schüler oder mit ihren Schülern sprachen, konnte zeigen,

dass etwa gleich viel positiver wie negativer Affekt gegenüber den Schülern sowohl im Sprachinhalt (Transkripte) als auch in der relativ gut zu kontrollierenden Gesichtsmimik zu finden war. Die kaum kontrollierbare Körpersprache dagegen strahlte bei Lehrern mit starker Stereotypneigung signifikant negativeren Affekt aus. Lehrer mit ausgeprägter stereotyper Tendenz zeigten insgesamt eine größere Diskrepanz zwischen intendierten und nicht intendierten affektiven Informationen. Auch wenn alle Lehrer sich bemühten, ihren Schülern warm und positiv zu begegnen, gelang dies nur den unvoreingenommenen Lehrern und diese unterschiedliche Haltung konnte bereits nach 10 Sekunden wahrgenommen werden.

Diese Resultate zeigen, dass Menschen sich in jeder Sekunde in irgendeiner Form nonverbal verhalten und sie damit ihr Innerstes ständig preisgeben. Oftmals reichen daher extrem kurze Beobachtungszeiträume, um seinen Gegenüber adäquat einschätzen zu können. Doch bisher bleibt ungeklärt, was in diesem Kontext unter Adäquatheit zu verstehen ist. Wann ist eine Einschätzung als genau, adäquat oder zutreffend anzusehen?

4.2.2. Indikatoren für Genauigkeit

Im Rahmen des „Thin slices“-Paradigmas kann eine Aussage über die Adäquatheit von intuitiven Einschätzungen über verschiedene Analysen erfolgen, die von der Art des bewerteten Merkmals abhängen.

Die untersuchten Eigenschaften lassen sich in zwei unterschiedliche Klassen von Merkmalen zusammenfassen. Zum einen gibt es Eigenschaften, die eine external messbare Entsprechung aufweisen wie beispielsweise Intelligenz. Menschen unterscheiden sich interindividuell anhand ihres Ausmaßes an Intelligenz und der jeweilige Ausprägungsgrad kann mit Hilfe objektiver Intelligenztests gemessen werden. Die Genauigkeit der Einschätzung wird in einem solchen Falle durch eine signifikante Korrelation zwischen der intuitiv zugesprochenen Intelligenz mit der tatsächlichen Ausprägung erzielt.

Die wenigsten im Interesse der psychologischen Forschung stehenden Merkmale weisen jedoch eine objektive oder gar physikalische Entsprechung in der Realität auf wie Größe oder Gewicht. In diesem Falle wird mit verschiedenen Möglichkeiten der Annäherung an Objektivität gearbeitet. Persönlichkeitseigenschaften beispielsweise werden mit speziell dafür entwickelten Tests gemessen. Möglich ist weiterhin, Selbsteinschätzungen, Einschätzungen von guten Bekannten, Familienmitgliedern, Kollegen, Vorgesetzten, Mitarbeitern, usw. miteinander zu vergleichen (Funder, 1999).

In diesem Falle wird Genauigkeit über den Indikator *Übereinstimmung* gemessen. Übereinstimmung ist definiert als Zusammenhang zwischen einer „Thin slices“- Beurteilung und einem Kriterium, wobei beide auf der gleichen Dimension gemessen werden (Carney, Colvin & Hall, 2007). Beispielweise demonstrierten Ambady, Hallahan und Connor (1999), dass die vermutete sexuelle Orientierung basierend auf 1- beziehungsweise 10-sekündigen Videosequenzen signifikant mit der von den Zielpersonen angegebenen sexuellen Orientierung übereinstimmte.

Im Gegensatz zu dieser Klasse von Eigenschaften stehen Merkmale mit vorwiegend sozialem Charakter. Diese werden erst dann zu einem „objektiven“ Attribut, wenn sich die Mehrheit der Personen in Bezug auf den Ausprägungsgrad des in Frage stehenden Attributs bei einer bestimmten Person einig ist (Funder, 1995). Ein typisches Beispiel für ein soziales Merkmal ist Höflichkeit. Ein Mensch ist erst dann höflich, wenn er als höflich angesehenes Verhalten zeigt und ihn daraufhin die Mehrheit seiner Mitmenschen als höflich wahrnimmt. In diesem Falle wird Genauigkeit entweder über interpersonellen Konsens oder über die Vorhersage von Verhalten gemessen.

Vorhersage als der andere mögliche Indikator für Genauigkeit ist definiert als Beziehung zwischen intuitiver „Thin slices“- Einschätzung und einem Verhalten oder einem Ergebnis, das einen anderen Gehalt als die Intuitivbeurteilung aufweist (Carney et al., 2007). So

konnten Ambady und Rosenthal (1993) beispielsweise zeigen, dass die Bewertung eines Lehrers durch seine Schüler am Ende des Schuljahres durch eine „Thin slices“- Bewertung ihres Enthusiasmus, ihrer Aufmerksamkeit und ihrer Wärme vorhergesagt werden konnte.

4.2.3.Kritische Anmerkungen und Bewertung des „Thin slices“- Paradigmas

Was ist nun zu lernen aus der weiter oben beschriebenen Fülle von Ergebnissen? Festzuhalten bleibt, dass Menschen scheinbar eine gewisse Kompetenz aufweisen, rasche und generell valide Einschätzungen anderer Menschen vorzunehmen, auch wenn sie diese kaum oder gar nicht kennen. Natürlich sind diese Eindrücke nicht gleichzusetzen mit absoluter Genauigkeit. Da sie jedoch zumindest minimal diagnostisch sind, können diese Ergebnisse als Untermauerung für die Hypothese des Persönlichkeitseinschätzungs-Instinktes gelten.

Ein wichtiger Beitrag der „Thin slices“- Methode liegt darin, dass ihre Ergebnisse auf Stimulusmaterial beruhen, welches die natürlich vorkommenden Varianz menschlicher Merkmale abbildet und damit der Wirklichkeit weitaus näher kommt als varianzanalytische Kontrastierungen auf Basis von statischen Fotos mit extremen Ausprägungen der zu bewertenden Merkmale (beispielsweise von sehr unattraktiven versus sehr attraktiven Personen), Gesichtsschemata, abstrakten Personenbeschreibungen, usw. Auch im tatsächlichen Leben beurteilt man täglich andere Menschen, die sich in irgendeiner Art und Weise verhalten, miteinander interagieren und sich unter Umständen, bezogen auf ihre Eigenschaften und Charakteristiken, nur in Nuancen voneinander unterscheiden.

Allerdings bleibt die Forschungsmethode nicht unkritisiert. Ein Teil der Kritik bezieht sich auf die Beurteilungssituation, die – auch wenn dem realen Leben näher – trotzdem eine Laborsituation darstellt. Wenig realistisch ist beispielsweise die unmittelbare Einschätzung

nach jeder Verhaltenssequenz. Anders als in realen Situationen wird hier der eingehende Informationsfluss nach einer definierten Zeit unterbrochen. So wird der Eindruck einzig auf Basis der bisher wahrgenommenen Informationen gebildet. Dies geschieht ohne dass gleichzeitig weitere Informationen auf den Organismus einströmen, die gleichsam verarbeitet und integriert werden müssen, so wie es in einer realen Situation der Fall wäre. Auch wenn dieser Eindruck eher auf ökologisch stimmigen Informationen beruht, so bleibt es eine realitätsferne Situation, die es dem ersten Eindruck relativ einfach macht, genau zu sein, da erstens der Strom eingehender Informationen gestoppt wird und zweitens die Beurteiler nicht mit den zu beurteilenden Personen interagieren.

Die Tatsache, dass die Zielpersonen nicht im Laufe einer tatsächlich stattfindenden Interaktion eingeschätzt werden, verändert die Situation des Beurteilenden entscheidend. Grundsätzlich betrachtet impliziert das Fehlen einer Interaktion gleichsam eine fehlende Verhaltensabstimmung. In realen Begegnungen beeinflussen sich die Akteure gegenseitig, das heißt jegliches Verhalten ist gleichzeitig sowohl eine Aktion als auch eine Reaktion. Sieht ein Beurteiler das Verhalten einer Zielperson in einer interaktionsfreien Situation, so wird dieses Verhalten mehr über die Person aussagen als Verhalten, das auf den Interaktionspartner abgestimmt und von ihm beeinflusst ist. Allerdings gibt es auch Studien, in denen die Zielpersonen in einer Interaktion gefilmt wurden (z.B. Babad et al., 1989; Vogt & Colvin, 2003). Da der Beurteiler jedoch kein Akteur in der Begegnung ist und somit das Geschehen weder beeinflusst noch von ihm beeinflusst wird, nimmt er auch in diesem Fall eine Außenperspektive ein. Diese erlaubt ihm, das Verhalten der Beteiligten aus einer anderen Perspektive zu betrachten, als dies in realen Situationen der Fall wäre.

Zwei weitere Kritikpunkte beziehen sich auf statistische Komponenten. Zum einen wird Genauigkeit nicht durch den Vergleich eines

„wahren“ Wertes mit der intuitiven Beurteilung für jedes einzelne Individuum berechnet, sondern durch eine Aggregation der Einzelurteile zu einem Gruppenurteil. Die einzelnen voneinander unabhängigen Einschätzungen über eine Wahrscheinlichkeit werden gemittelt und mit dem jeweiligen Kriterium verglichen. Das kollektive Urteil gibt somit nicht die Meinung eines einzelnen Intuitivbeurteilers wieder, sondern bringt zum Ausdruck, was alle Beurteiler im Durchschnitt denken (Surowiecki, 2005).

Jedes Einzelurteil beinhaltet generell zwei Komponenten: Information, das heißt eine akkurate Einschätzung des wahren Wertes der Zielperson, und Irrtum, resultierend aus spezifischen Verzerrungen und Wahrnehmungsfehlern des Individuums. Wird der Durchschnitt aus den einzelnen Einschätzungen vieler Intuitivbeurteiler berechnet, so wird der Irrtum subtrahiert und es bleibt die Information. Die Fehler der Einzelnen gleichen sich gegenseitig aus (Surowiecki, 2005). Die Gruppe gibt somit immer bessere und verlässlichere Urteile und Prognosen ab als ein einzelner, kompetenter Entscheidungsträger. Folglich ist die Validität jedes Einzelurteils weitaus geringer als die Genauigkeit des Gruppenurteils (Kunda, 1999). In diesem Zusammenhang spricht man auch von der „Weisheit der Vielen“ oder dem „Wisdom of crowds“-Effekt (Surowiecki, 2005).

Der Kerngedanke dieser „Weisheit der Vielen“ besteht allerdings nicht darin, dass eine Gruppe stets zu einem vollständig korrekten Urteil kommt, sondern im Schnitt mit einem valideren Urteil aufwartet als ein Einzelner dies vermag. Zu vermerken bleibt, dass die Aggregation die Validität der Einschätzung steigert und somit die Korrelation im Anschluss vergrößert. Folglich wird keine Aussage darüber gemacht, wie gut der Einzelne in einer realen Situation eine andere Person hinsichtlich des gegebenen Merkmals beurteilen könnte.

Um nach der Aggregation der Einzelwerte eine Aussage über die Genauigkeit der Personenwahrnehmung abzugeben, wird in der Regel ein statistischer Test gegen die Null-Hypothese beziehungsweise

gegen Zufallseinschätzungen durchgeführt. Doch was genau sagt eine statistische Signifikanz in diesem Fall aus?

Auf diese Frage bezieht sich die zweite Komponente der statistischen Kritik. Auch wenn der gefundene Wert signifikant von einem zufällig zu erwartenden Ergebnis abweicht, kann der Effekt dennoch klein sein. Beispielsweise fanden Zebrowitz, Hall, Murphy und Rhodes (2002), dass Intelligenz akkurat aus der Mimik einer Person erkennbar ist. Der Effekt ist allerdings mit $r = .19$ lediglich als moderater Zusammenhang zu bewerten. Innerhalb des „Thin slices“-Paradigmas sind jedoch allgemeine Aussagen über Genauigkeit eher die Regel statt genaue Unterscheidungen in kleinere oder größere Effekte (Carney et al., 2007).

Abschließend bleibt festzuhalten, dass die „Thin slices“-Forschung erstaunliche Fähigkeiten des Menschen aufdeckt, viele Fragen jedoch bis heute unbeantwortet lässt. Wie genau bilden sich intuitive Eindrücke von anderen Personen? Welche Umstände oder Informationen beeinflussen diese Blitzurteile in welcher Weise? Wie sehr determinieren erste Eindrücke anschließende Denkprozesse und Folgeverhalten? Welche Informationen werden eigentlich über das nonverbale Verhalten der Zielpersonen vermittelt? Und welche Hinweise werden von den Intuitivbeurteilern für die Einschätzung bestimmter Eigenschaften herangezogen?

Auf keine dieser Fragen liefert die „Thin slices“-Forschung eine Antwort. So wie die kognitive Sozialpsychologie nur den Prozess der Eindrucksbildung betrachtet und dabei die Frage nach der Genauigkeit weitgehend ignoriert, prüft das „Thin slices“-Paradigma nur das Ergebnis von Personenwahrnehmung und vernachlässigt dabei nahezu vollständig den Prozess. Mit anderen Worten ausgedrückt widersprechen die Ergebnisse des „Thin-slices“-Paradigmas den Ergebnissen der sozialen Kognitionsforschung nicht, da schlichtweg andere Fragen gestellt werden.

In diesem Punkt erweitern die vorliegenden Untersuchungen den bisherigen Forschungsstand, da sie an der Schnittstelle zwischen beiden Paradigmen angelegt sind.

Resultierend aus den bisherigen Ergebnissen der „Thin slices“-Forschung wird die Hypothese aufgestellt, dass die intuitiven Beurteiler in der Lage sind, die tatsächliche Führungskraft der Zielpersonen besser als zufällig zu erkennen.

Doch neben der Genauigkeit der Personenwahrnehmung ist auch die Untersuchung des Beurteilungsprozesses Gegenstand aller vier Studien. Insofern ist im vorliegenden Fall die Frage wichtig, wie die intuitiven Beurteiler zu ihrer Einschätzung gelangen, auch wenn diesbezüglich aus der „Thin slices“-Forschung keine Hypothese abgeleitet werden kann.

Eine mögliche Beleuchtung des Urteilsprozesses aus der evolutionären Perspektive kann aus dem Ansatz der „fast and frugal heuristics“ nach Gigerenzer und Kollegen heraus stattfinden, der im Folgenden vorgestellt wird.

4.3. Eine Integration: schnelle und sparsame Heuristiken

Heuristiken, so haben wir bisher gelernt, erlauben eine schnelle soziale Orientierung, vernachlässigen dabei jedoch wichtige Informationen, führen manchmal fundamental in die Irre und verstoßen oftmals gegen die Gesetze der Logik. Auf der anderen Seite konnte eine Reihe von Studien im Rahmen der „Thin slices“-Forschung zeigen, dass Menschen recht gute Schlüsse in Bezug auf Charakter, Emotionen und zukünftigem Verhalten unbekannter Personen zu ziehen vermögen, obwohl sie deren Verhalten nur einen Moment in einer ganz spezifischen Situation beobachten konnten. Wie passen diese sich augenscheinlich widersprechenden Befunde zusammen? Könnte es vielleicht doch sein, dass intuitive „Abkürzungen“ wie

Heuristiken und Stereotype zum gewünschten Ziel führen, ohne dabei schwerwiegende Fehler zu begehen?

Stimmt man dieser Aussage zu, so befindet man sich in einer Linie mit der Position von Gerd Gigerenzer und der ABC Forschergruppe (Gigerenzer, 2002a, 2002b, 2007, 1996; Gigerenzer et al., 2001; Todd & Gigerenzer, 2003). Ausgangspunkt dieser Forschung ist ihre Kritik an den Annahmen und Methoden der klassischen kognitiven Sozialpsychologie und deren Fokus auf Irrtümer und Fehler im Denk- und Entscheidungsprozess. Eines ihrer Argumente gegen die soziale Kognitionsforschung lautet, dass viele der gefundenen Fehler verschwinden, wenn man den Inhalt oder den Kontext der Aufgabe für die Probanden verändere.

So konnte gezeigt werden, dass Menschen bestimmte Aufgaben deutlich besser lösen können, wenn die Aufgabe realen Problemstellungen näher kommt. Ein Beispiel in diesem Kontext ist die Angabe von absoluten Zahlen statt Prozentsen bei Schätzaufgaben, da Menschen im Umgang mit natürlichen Häufigkeiten geübter sind (Gigerenzer, 2002a, 2002b, 2007). Diese Ergebnisse werden von der sozialen Kognitionsforschung zumeist ignoriert, da ihrer Auffassung nach bei solchen Aufgaben nur die Form, nicht aber der Inhalt zählt. Da im Fall von Prozentsen versus Zahlen beide Aufgaben mathematisch und logisch die gleichen sind, sollte die Darstellung keine Rolle spielen. Beweise dafür, dass die Präsentation einen Einfluss auf die Lösungswahrscheinlichkeit nimmt, wird als weiterer Indikator für menschliche Irrationalität gewertet (Gigerenzer, 2002b).

Im Zentrum der Kritik steht jedoch das Fehlen präziser und damit falsifizierbarer Modellierungen von Denkprozessen. Durch eng gefasste Zielkriterien und normativ als korrekt angenommene Inferenzprozesse gelingt es leicht nachzuweisen, dass Menschen in ihrem Denken oftmals vom rationalen Pfad der Logik abweichen (Gigerenzer, 1996).

Im Gegensatz zur klassischen Kognitionsforschung betrachtet die ABC Forschergemeinschaft soziale Rationalität (Todd & Gigerenzer, 2003). Menschliches Denken und Entscheiden wird unter dem Gesichtspunkt der sozialen Struktur der Umgebung betrachtet. Ausgehend von dieser Prämisse erweisen sich Prozesse und Resultate, die im Labor als Fehler oder Verzerrung definiert würden, im realen Leben als durchaus sinnvolle Denk- und Entscheidungsweisen (Gigerenzer, 2002a).

Wenn es um die Einschätzung oder den Umgang mit anderen Personen geht, folgen Menschen ihrem sozialen Instinkt. Soziale Instinkte stehen dem Menschen natürlich zur Verfügung und ermöglichen schnelle und einfache Lösungen für komplexe soziale Probleme. Dazu Gigerenzers eigene Worten: „Doch der gewöhnliche Homo sapiens ist ein soziales Tier mit der Fähigkeit, rasche Schlussfolgerungen über die Dynamik des sozialen Lebens zu ziehen und Gesichtspunkte wie Verrat, Vertrauen und Ansehen zu berücksichtigen. Wir haben nicht nur die Fähigkeit über die vorliegende Information hinauszugehen, sondern sind auch nicht in der Lage, diese Fähigkeit nicht zu verwenden. Wir können nicht aufhören, Schlüsse über andere zu ziehen.“ (Gigerenzer, 2007, S. 221)

Diese Schlüsse manifestieren sich in Bauchgefühlen, Intuitionen oder Ahnungen. Alle drei Begriffe beschreiben Urteile, die rasch im Bewusstsein auftauchen, deren tiefere Gründe nicht ganz bewusst sind und die stark genug sind, damit ein Mensch danach handelt. Die soziale Intelligenz besteht im Prinzip aus einfachen Faustregeln, Heuristiken, die schnell und ökonomisch arbeiten. Um ein Problem zu lösen greifen Heuristiken die wichtigsten Informationen heraus und lassen gleichzeitig alle anderen außer Acht. Sie machen sich die evolvierten Fähigkeiten des Gehirns zunutze, um gute soziale Entscheidungen zu fällen. Zu diesen Fähigkeiten gehören unter anderem Vertrauen, Nachahmung, Wiedererkennen oder Liebe. Wären Menschen beispielsweise unfähig, Empathie zu empfinden, so könnte

kein Mensch, egal in welcher Situation und aufgrund welchen Verhaltens, auf die Wünsche und Absichten seines Gegenübers schließen (Gigerenzer, 2007).

Die Struktur der sozialen Umwelt bestimmt dabei, wie gut oder schlecht eine Heuristik funktioniert. Die wahre Intelligenz von Bauchgefühlen enthüllt sich darin, dass durch eine schnelle Evaluation der Situation entschieden wird, welche Faustregel für die gegebene Situation die angemessene ist. Wie dem Handwerker für jedes Problem ein spezielles Werkzeug zur Verfügung steht, so bedient sich die soziale Intelligenz aus einem Werkzeugkasten voller Daumenregeln, die genetisch, kulturell oder individuell entwickelt oder übermittelt wurden.

Beispielsweise lesen Menschen im sozialen Miteinander sehr viel aus den Blicken ihrer Mitmenschen ab. Diese Blickheuristik macht sich die evolvierte Fähigkeit zunutze, Objekte mit den Augen verfolgen zu können. Neuankömmlinge in einer Gruppe von Menschen (oder hierarchisch organisierten Primaten) können den Sozialstatus der Gruppenmitglieder dadurch erkennen, wer wen anblickt (Cosmides & Tooby, 1992). Eine sorgfältige Blickverfolgung kann somit der Vermeidung von Konflikten oder der Eingliederung in eine bestehende Hierarchie dienen. Schon im Alter von etwa zwei Jahren beginnen Kinder, die Blicke und die Körperbewegungen anderer zu beobachten, um diese als Hinweis auf deren Wünsche und Absichten zu nutzen. Im Alter von drei Jahren lernen Kinder dann, bestimmte Blicke als Hinweise auf Täuschungen zu verstehen (Freire, Eskritt & Lee, 2004).

Die Blickheuristik, das heißt die Verfolgung des Blickes und die über diese Information hinausgehenden Schlussfolgerungen, stellt somit eine einfache, auf einer evolvierten Fähigkeit aufbauende Faustregel dar. Diese ist trotz ihrer Einfachheit offensichtlich in der Lage, ein breites Spektrum komplexer sozialer Probleme zu lösen. Sie ist jedoch nicht in allen sozialen Situationen die richtige Daumenregel.

Die Wahl und die Anwendung der richtigen Heuristik für eine spezielle Situation sind die Fähigkeiten, die einen Menschen zu sozial intelligenten Bewertungen und Entscheidungen befähigen. Die meisten sozialen Interaktionen sind laut Gigerenzer (2007) weniger ein Ergebnis komplexer Kalkulationen, sondern vielmehr das Resultat von sozialen Instinkten oder Ein-Grund-Entscheidungen. Zwar handelt es sich bei sozialen Entscheidungen immer um Entscheidungen unter Ungewissheit, jedoch muss nach Gigerenzers Auffassung im Urteilsprozess eine gute Intuition immer Informationen außer Acht lassen.

Ein-Grund-Entscheidungen sind weiter verbreitet, als man zunächst vermuten mag. Sexuelle Selektion beispielsweise beruht oftmals auf einem einzigen Signal. Als prototypisches Exempel für sexuelle Selektion wird oftmals die Partnerwahl bei Pfauen herangezogen. Je mehr Augenflecken das Gefieder eines männlichen Pfaus aufweist, desto größer ist sein Paarungserfolg (Petrie & Halliday, 1994). Pfauenhennen entscheiden sich somit aus einem einzigen Grund für einen Partner: dem Gefieder. Doch ist das wirklich eine sinnvolle Strategie? Wäre eine elaboriertere Partnerwahl, eine, auf mehreren Merkmalen basiert, nicht eine bessere?

Intuitiv würde man diese Frage wahrscheinlich bejahen. Die biologische Forschung zeigt jedoch, dass die Schwanzfedern eines Pfaus tatsächlich einen sehr guten Hinweis auf seine Gesundheit, Kraft, Überlebensfähigkeit und genetische Qualitäten darstellen (Gadagkar, 2003). Pfauenhennen tun also gut daran, sich an diesem einen Merkmal zu orientieren.

Ein menschlicheres Beispiel findet sich im Wahlverhalten deutscher Wähler. Neben aller Komplexität menschlichen Wissens, Informationsverarbeitungsstrukturen und Einstellungen scheint die Ansiedlung einer Partei auf dem Links-Rechts-Kontinuum politischer Einstellungen in über 90 Prozent der Wahlentscheidungen der ausschlaggebende Grund zu sein (Gigerenzer, 1982). Insgesamt führten Gigerenzer und Kollegen eine ganze Reihe von unterschiedli-

chen Studien durch, die belegen, dass sich Menschen bei ihren intuitiven Urteilen vielfach lediglich von einem guten Grund leiten lassen (Bröder, 2003; Gigerenzer et al., 2001).

Nun reicht jedoch nicht in allen Situationen ein einziger Grund als Entscheidungsgrundlage aus. Ein Prozess, bei dem man nacheinander so viele Gründe heranzieht, bis einer von ihnen eine Entscheidung zulässt, bezeichnet man als sequentielle Entscheidungsfindung. Bei diesem lexikografischen Prozess werden mehrere Gründe in der Reihenfolge ihrer Bedeutung aus dem Gedächtnis abgerufen und geprüft. Sobald sich zwei zur Auswahl stehende Alternativen anhand eines Grundes unterscheiden, tritt die Stopppregel in Kraft. Besteht somit ein Überhang zu einer Alternative, wird der Prozess abgebrochen und das Urteil gefällt, schlussendlich wieder nur aus einem einzigen Grund.

So könnte ein Personalverantwortlicher, der für die Besetzung einer vakanten Führungsposition die Auswahl zwischen zwei Kandidaten hat, zunächst die Anzahl der Jahre vergleichen, die die beiden Bewerber bereits in einer ähnlichen Position gearbeitet haben. Führt dies zu einer Differenzierung der beiden Kandidaten, so könnte er sich für den Kandidaten mit der längeren Berufserfahrung entscheiden und würde somit einer Ein-Grund-Heuristik folgen. Hätten beide Kandidaten die gleiche Berufserfahrung, so würde dieser Anhaltspunkt für eine Entscheidung nicht ausreichen und das nächste Kriterium müsste herangezogen werden, beispielsweise die Branche, in der die Bewerber bisher eine Führungsposition bekleidet haben. So könnte der Personalreferent ein Merkmal nach dem anderen prüfen, bis das erste einen Unterschied zwischen den beiden Kandidaten aufzeigt. Daraufhin würde er sich für den Bewerber entscheiden, der, gemessen an dieser Eigenschaft, im direkten Vergleich besser abschneidet.

Vergleicht man verschiedene Ein-Grund-Entscheidungen mit Urteilen, die auf Basis einer komplexen Algebra gefällt wurden, so zeigt

sich, dass erstere – gemessen an den Kriterien Genauigkeit und Sparsamkeit - nicht schlechter abschneiden als letztere (Gigerenzer & Goldstein, 2001). Wenn Versuchspersonen schätzen sollen, welches von zwei Objekten größer, effektiver, gefährlicher usw. ist, so führt eine Take-the-Best-Heuristik, das heißt folge dem einen, wichtigsten Grund, zu besseren Vorhersagen und das, obwohl sie gleichzeitig mit weniger Informationen auskommt. Für das obige Beispiel würde dies bedeuten, dass ein Personalmanager, der nur nach der Anzahl der Berufsjahre in einer vergleichbaren Position prüft (vorausgesetzt dieser Anhaltspunkt ist der beste zur Voraussage des zukünftigen Berufserfolges), eine schnellere aber genauso gute Entscheidung fällt wie sein Kollege, der jeweils Berufsjahre, Alter, Schul- und Hochschulabschluss, bisheriges Gehalt, Branche, etc., genau gewichtet und gegeneinander abwägt. Dieses Schema konnte von Gigerenzer und Kollegen anhand von über zwanzig realen Problemen aus Psychologie, Wirtschaft, Biologie, Soziologie, Gesundheitswesen und anderen Bereichen repliziert werden (Gigerenzer & Goldstein, 2001).

Intuitionen, die auf einem guten Grund beruhen, sind nicht nur ökonomisch, sondern oftmals auch außerordentlich zutreffend, weil sie die nützlichsten Informationen herausgreifen und alle anderen außer Acht lassen. Genau wie innerhalb der sozialen Kognitionsforschung wird auch nach diesem Ansatz angenommen, dass Menschen Heuristiken bei der Urteilsbildung verwenden. Der entscheidende Unterschied besteht darin, dass hier diese Nutzung im Durchschnitt gute Ergebnisse produzieren soll.

Gigerenzers Fazit lautet, dass Menschen gut beraten sind, sich auf ihre Intuitionen zu verlassen. Dies gilt insbesondere wenn sie über wenig relevante Informationen verfügen, einen unbekannten gegenwärtigen Zustand voraussagen sollen und im Allgemeinen für schwer vorhersagbare Situationen. All dies ist in der vorliegenden Studie der Fall. Die Beobachter sehen die Zielpersonen nur sehr kurz, ohne Ton, alleine und in einer kontextarmen Situation. Der Schluss von einer

nichtsozialen Situation auf ein soziales Merkmal und die Vorhersage von sozialem Verhalten ist eine schwierige Aufgabe, die auf einem Minimum an Informationen beruhen muss.

Auch wenn bisher keine Studie die Relevanz von Ein-Grund-Heuristiken im Bereich Personenwahrnehmung untersucht hat, liegt die Vermutung nahe, dass Beurteiler simple und sparsame Heuristiken verwenden, das heißt sich an wenigen Merkmalen orientieren, um führungsstarke von führungsschwachen Personen zu unterscheiden. Darüber hinaus wird für die vorliegende Studienreihe angenommen, dass Menschen die Führungsstärke anderer zutreffend erkennen können.

Eine zutreffende Wahrnehmung von Führungsstärke fremder Personen unter Rückgriff auf simple und sparsame Heuristiken kann nur gelingen, wenn die Heuristik selber beziehungsweise der gewählte Orientierungspunkt eine sinnvolle Herangehensweise darstellt. Stellt man Vermutungen auf über personenbezogene Merkmale, an denen sich die Beurteiler bei der Zuschreibung von Führungsstärke orientieren könnten, so liegen auch aus der evolutionären Perspektive heraus das Geschlecht und die Attraktivität der Zielpersonen nahe. Die Nutzung dieser Heuristiken kann dann zu einem zutreffenden Urteil führen, wenn diese Merkmale valide Indikatoren für Führungsstärke darstellen, wenn sie also ein „Körnchen Wahrheit“ besitzen.

4.3.1. Die Hypothese des „Körnchen Wahrheit“

Die „Kernel of truth“- oder „Körnchen Wahrheit“- Hypothese geht davon aus, dass Stereotype in der sozialen Wahrnehmung einen wahrheitsgetreuen Kern besitzen, da diese von tatsächlichen Charakteristiken einer Gruppe beeinflusst werden (Prothro & Melikian, 1955). Diese Linie der Stereotypenforschung stellt eine Ausnahme in der klassischen Stereotypforschung dar, da sie sich mit der Frage nach der Exaktheit von Stereotypen befasst und nicht mit dem Stereotypisierungsprozess.

Ergebnisse dieses Forschungsansatzes belegen einerseits, dass einige Stereotype vollkommen unzutreffend sind (Pettigrew, 1964). Andererseits zeigte sich in anderen Studien tatsächlich ein „Körnchen Wahrheit“, beispielsweise in den stereotypen Vorstellungen die verschiedene Volksgruppen übereinander haben (Buchanan & Cantril, 1953). Auch das Stereotyp der unterschiedlichen Körpergröße von Männern und Frauen erwies sich als zutreffend. Dies mag zunächst trivial erscheinen, doch werden auch Entwicklungsphasen, in denen Mädchen größer sind als Jungen ihres Alters, korrekt getroffen (Biernat, 1993).

Als Ursprung dieses „Körnchen Wahrheits“ in der sozialen Wahrnehmung wird einerseits die Möglichkeit diskutiert, dass bestimmte geteilte Gene zu einer ähnlichen Ausprägung bestimmter Persönlichkeitszüge führen (Penton-Voak, Pound, Little & Perrett, 2006).

Alternativ könnte der wahrheitsgemäße Kern einiger Stereotype auch das Resultat einer Sich-selbst-erfüllenden-Prophezeiung sein (Gilovich, Keltner & Nisbett, 2006).

Gehen wir also davon aus, Männern und attraktiveren Personen werde eine höhere Führungsstärke zugeschrieben. Könnten diese Merkmale ebenfalls ein „Körnchen Wahrheit“ besitzen und wenn ja, wie könnte man es aus der evolutionären Vergangenheit heraus erklären?

Führung entwickelte sich vermutlich aus verschiedenen Herausforderungen an Gruppen und der daraus resultierenden Notwendigkeit der Gruppenkoordinierung (Van Vugt, Hogan & Kaiser, 2008). Zu diesen Herausforderungen zählten beispielsweise die Suche nach neuen Lagerstätten, Migrationswanderungen und Konflikte innerhalb einer sowie zwischen verschiedenen Gruppen. Zur Lösung dieser Koordinierungsprobleme kam es zur Bildung verschiedener Rollen innerhalb von Gruppen. Dabei wurden die Rollen der Führenden, der Personen mit dem größten Einfluss in der Gruppe und dem höchsten sozialen Status, traditionellerweise von den besten Jägern bezie-

hungsweise den besten Kämpfern gehalten, den „Big Men“ (Boehm, 1999). Dieses Muster hielt sich über Jahrtausende, auch nach den massiven Veränderungen der Lebensumstände durch den Beginn der Landwirtschaft vor etwa 13.000 Jahren. Mit steigender Gruppengröße und Ansammlung von Gütern, stieg das Potenzial für Intra- und Intergruppenkonflikte. Kriegsherren, sowie aggressive und durchsetzungsstarke Führer erlangten die höchsten Positionen innerhalb der damaligen Gesellschaften (Johnson & Earle, 2000). Somit kann festgehalten werden, dass Führungspersonen innerhalb einer Gruppe in der Regel jeweils diejenigen waren, die sich durch bestimmte Fähigkeiten wie Jagd- oder Kampffähigkeiten, Stärke und eine dominante Strategie gegen die anderen Konkurrenten um diese Position erfolgreich durchgesetzt hatten.

Betrachtet man zusätzlich das große Risiko, das aus der Rolle des Gefolgsmannes resultiert, sollten Gruppen weiterhin Führende bevorzugen, die dauerhaft dem Wohl der Gruppe zuträglich sind (Van Vugt et al., 2008).

Diese beiden Annahmen zusammen erklären die bis heute nachweisbaren Zusammenhänge von Führung mit Alter, Größe und Gesundheit. Dabei äußert sich Gesundheit unter anderem durch physische Attraktivität. Die Forschung zeigt, dass gerade die äußerlichen Merkmale allgemein als attraktiv wahrgenommen werden, die Hinweise auf Gesundheit, hohe Krankheitsresistenz und genetische Qualität liefern. Dazu gehören beispielsweise das Taille-Hüfte-Verhältnis, Symmetrie des Gesichts sowie Körpersymmetrie (für einen Überblick siehe Berry, 2000).

Aus dieser langen Tradition vornehmlich männlicher und physisch herausragender Führungspersonen sollten sich psychologische Mechanismen zur Bevorzugung von attraktiven (männlichen) Führungspersonen entwickelt haben. Dieser evolutionären Vergangenheit mag ebenso die bis heute anhaltende maskuline Typisierung von Führungspositionen entstammen und damit einhergehend könnte das

Resultat eine höhere wahrgenommene Führungsstärke bei Männern sein sowie eine Assoziation von Führungsstärke und Attraktivität.

Betrachtet man das Gesamtbild der evolutionären Annahmen und Ergebnisse, so lässt sich die Hypothese aufstellen, dass die intuitiven Beurteiler die Merkmale Geschlecht und Attraktivität als simple und sparsame Heuristiken benutzen. Diese könnten zu einer zutreffenden Wahrnehmung von Führungsstärke führen, wenn beide ein „Körnchen Wahrheit“ besäßen.

4.4. Der „sexual attribution bias“

Eine andere Forschungsreihe innerhalb der evolutionären Psychologie beschäftigt sich mit den Merkmalen Geschlecht und Attraktivität, beziehungsweise deren Zusammenwirken, im Kontext der sexuellen Selektion. Eine Reihe von Studien in diesem Rahmen konnte zeigen, dass es bei der Einschätzungen fremder Personen je nach Attraktivität und Geschlecht der zu beurteilenden Personen und dem Geschlecht des Beurteilenden zu einem speziellen Attributionsmuster kommt, welches „sexual attribution bias“ genannt wurde (Försterling, Preikschas & Agthe, 2007).

Beim „sexual attribution bias“ wirken die Merkmale Geschlecht und Attraktivität auf eine spezielle Art und Weise zusammen, so dass es zu einem spezifischen Präferenzmuster je nach Geschlecht des Beurteilenden kommt. Dabei sind zwei Komponenten zu unterscheiden, die Abwertungskomponente (*derogation component*) und die Glorifizierungskomponente (*glorification component*). Letztere äußert sich darin, dass attraktiven gegengeschlechtlichen Personen insgesamt weitere positivere Merkmale zugeschrieben werden, während die Abwertungskomponente besagt, dass attraktiven gleichgeschlechtlichen Personen weniger positive Eigenschaften zugesprochen werden (Försterling et al., 2007). Doch wie kommt dies zustande?

Das spezielle Zusammenspiel von Attraktivitäts- und Geschlechterstereotyp basiert nach Meinung der Autoren auf evolvierten psychologischen Mechanismen. Um diese Annahme zu verstehen, gilt es, die Funktion der beiden einzelnen Komponenten zu betrachten. Die Attraktivität von Personen ist im evolutionären Kontext von großer Bedeutung, da die Charakteristiken, die einen Menschen attraktiv für das andere Geschlecht machen, bei Frauen Hinweise auf ihre Fruchtbarkeit und ihren reproduktiven Wert und allgemein einen Hinweis auf genetische Qualität liefern (z.B. Berry, 2000, Sugiyama, 2005). Da Attraktivität somit im Kontext der eigenen Reproduktion, die das ultimate evolutionäre Ziel darstellt, besonders relevant ist, sollten Menschen allgemein außerordentlich sensibel auf diesen Hinweis reagieren.

Bei der Partnerwahl und der Partnerbindung schreiben Försterling und Kollegen (2007) dem Glorifizierungsfaktor eine wichtige evolutionäre Funktion zu. Schreibt beispielsweise eine Frau einem körperlich attraktiven potenziellen Partner noch weitere gute Eigenschaften zu, so verstärkt dies dessen Anziehungskraft auf sie und die Wahrscheinlichkeit erhöht sich, dass sie einerseits subjektiv passende und begehrten Partner trifft und sich andererseits auch um eben diese bemüht. Weiterhin führt eine glorifizierende Wahrnehmung innerhalb einer Partnerschaft dazu, dass die Beziehung eher aufrecht erhalten wird (Kelley, 1979). Evolutionär gesprochen sind die Kosten geringer als der Nutzen, wenn gegengeschlechtliche Personen verherrlicht wahrgenommen werden, da keine Konkurrenz um Ressourcen besteht.

Die Abwertungskomponente dagegen kommt im intrasexuellen Wettbewerb zum Tragen, das heißt im Wettbewerb innerhalb eines Geschlechts um den Zugang zu erstrebenswerten Partnern. Die Herabsetzung attraktiver gleichgeschlechtlicher Konkurrenten hat intraindividuell den Vorteil, dass das eigene Selbstwertgefühl und das Durchhaltevermögen unberührt bleiben. Auf sozialer Ebene führt

eine nach außen kommunizierte Abwertung vermutlich zu einer Verringerung der Wettbewerbschancen des gleichgeschlechtlichen Konkurrenten. Beide Aspekte führen zu einem Ziel: die Stärkung der eigenen Position im intrasexuellen Wettbewerb.

In einer Reihe von Studien konnten Luxen und Van de Vijver (2006) zeigen, dass in simulierten Bewerbungssituationen studentische Entscheider je nach ihrem Geschlecht spezifische Präferenzmuster hinsichtlich der potenziellen Kandidaten zeigten. Wenn die erwartete Interaktionshäufigkeit mit dem ausgewählten „Bewerber“ hoch war, zeigten beide Geschlechter gegengeschlechtlich eine Präferenz für attraktive Personen. Weibliche Versuchsteilnehmer gaben dagegen bei gleichgeschlechtlichen Auswahlentscheidungen den unattraktiveren Kandidatinnen im Vergleich mit attraktiven Kandidatinnen den Vorzug. In einer Folgestudie mit Personalreferenten konnten die Ergebnisse nur teilweise bestätigt werden. Der direkte Einfluss der Attraktivitätsheuristik konnte repliziert werden, jedoch zeigten Frauen diesmal keine Abwertung attraktiver Kandidatinnen.

Die Studien von Agthe, Spörrle, Dörfler und Försterling (2008) gingen der Vermutung nach, dass diese Attributions- und Bevorzugungsmuster auch einen Einfluss auf tatsächliche Auswahlentscheidungen haben sollten. Dabei wurde sowohl das Stimulusmaterial als auch die Auswahlentscheidung systematisch variiert. Als Stimulusmaterial wurden entweder Fotos oder Videosequenzen dargeboten und bei den Auswahlentscheidungen handelte es sich entweder um die Auswahl eines Kandidaten für eine Arbeitsstelle, zur Vergabe eines Stipendiums oder für die Vergabe eines Studienplatzes.

In diesen Studien wiesen Beurteiler eine allgemeine Präferenz für attraktive im Vergleich zu unattraktiven gegengeschlechtlichen Personen auf. Gleichzeitig wurden attraktive gleichgeschlechtliche Personen im Vergleich zu unattraktiven gleichgeschlechtlichen Personen eher abgewertet. Insgesamt zeigte sich, dass gleichgeschlechtlich vor allem mittelattraktive Personen bevorzugt wurden. Dieses Muster

konnte sowohl für auf die Zielperson gerichtete positive Emotionen als auch für angestrebte soziale Interaktionen gefunden werden. Mit anderen Worten ausgedrückt besagt dieses Ergebnis, dass Frauen sich lieber mit unattraktiven bis mittelattraktiven Frauen und mit attraktiven Männern umgeben, während Männer diesbezüglich attraktiven Frauen sowie unattraktiven bis mittelattraktiven Männern den Vorzug geben.

In allen durchgeführten Studien zeigte sich eine Dreifachinteraktion aus Geschlecht und Attraktivität des Beurteilten sowie dem Geschlecht des Beurteilers. Auf gleichgeschlechtlicher Ebene wurden mittelattraktive Kandidaten bevorzugt, während gegengeschlechtlich jeweils der attraktivere Gegenpart bevorzugt wurde.

In vorhergehenden Studien konnte die Forschergruppe zeigen, dass weibliche und männliche Beurteiler den beruflichen Erfolg von zu beurteilenden fremden Personen je nach ihrem Geschlecht und ihrer Attraktivität unterschiedlich attribuierten. Der Erfolg von gleichgeschlechtlichen unattraktiven Stimuluspersonen wurde eher auf Fähigkeit attribuiert, während Erfolg von gleichgeschlechtlichen attraktiven Stimuluspersonen eher durch Glück erklärt wurde. Bei andersgeschlechtlichen Stimuluspersonen zeigt sich dagegen, dass der berufliche Erfolg mit steigender Attraktivität eher mit Fähigkeitsfaktoren begründet wurde (Försterling et al., 2007).

Ausgehend von diesen Überlegungen wird erwartet, dass auch die Wahrnehmung von Führungsstärke diesen Prozessen unterworfen ist. Doch wie genau würde er sich äußern?

Um von einem „sexual attribution bias“ sprechen zu können, müsste gegengeschlechtlich ein linearer Zusammenhang zwischen Attraktivität und der zugeschriebenen Führungsstärke zu finden sein. Je attraktiver eine Zielperson des anderen Geschlechts ist, desto mehr Führungsstärke wird ihr zugesprochen (Glorifizierungs-Komponente). Gleichzeitig müsste bei der gegengeschlechtlichen Einschätzung ein kurvilinearere Zusammenhang auftreten, in dem Sinne, dass mittel-

attraktive Zielpersonen bei der Zuschreibung von Führungsstärke begünstigt beurteilt werden (Abwertungs-Komponente).

Ebenso wie die soziale Kognitionsforschung befasst sich die Forschung zum „sexual attribution bias“ bislang vornehmlich mit dem Wahrnehmungsprozess, nicht jedoch mit dem Ergebnis dieses Prozesses. Führt man die Annahmen des „sexual attribution bias“ jedoch zu Ende, so gelangt man unweigerlich zu der Überlegung, dass unter Berufung auf diese beiden Merkmale kaum ein valides Urteil über die Führungsstärke fremder Personen gelingen kann, wenn dieses tatsächlich vor allem der Glorifizierung des anderen Geschlechts beziehungsweise der Abwertung gleichgeschlechtlicher Konkurrenten dient.

Die Resultate des „sexual attribution bias“ legen nahe, dass die menschliche Wahrnehmung dahingehend konstruiert ist, vorteilhaftes Handeln zu induzieren, insofern als dass sie Entscheidungen und Maßnahmen fördert, sich mit attraktiven potenziellen Partnern und mittelattraktiven Konkurrenten zu umgeben. Insofern könnte diese Wahrnehmung und das daraus resultierende Verhalten für die eigene Fitness von Nutzen sein, auch wenn sie nicht hilfreich für das Erkennen der wahren Eigenschaften anderer Personen ist.

5. Integrative Betrachtung und Schlussfolgerungen für die vorliegende Studie

In den 70er Jahren begann eine Revolution innerhalb der klassischen Sozialpsychologie (Kenrick et al., 2006). Mehr und mehr Forscher beschäftigten sich mit dem Einfluss kognitiver Prozesse auf soziales Verhalten, ein bis dato undenkbares Unterfangen. Nach anfänglich immenser Kritik, entwickelte sich die kognitive Sozialpsychologie schnell zu einem unentbehrlichen Zweig der Psychologie. Eine Seite der kognitiven Revolution war die stärkere Betonung von Prozessen, seien es Denk-, Gefühls-, Aktivierungs- oder Handlungsprozesse. Die

Genauigkeit sozialer Wahrnehmung wurde dagegen in diesem Zuge zur Randerscheinung degradiert (z.B. Krueger & Funder, 2004).

Eine andere Seite war die Fokussierung auf Fehler und Verzerrungen im Denk- und Entscheidungsprozess. Innerhalb dieses Paradigmas ist die Richtschnur die rein mathematische Logik. Ihre Entsprechung entscheidet darüber, ob ein Urteil als rational oder als irrational gilt. Doch oftmals scheitern Menschen an der Aufgabe, korrekte Lösungen für wohl definierte Probleme zu finden. Der Schwerpunkt auf Prozesse geht jedoch zu Lasten der Frage nach den Inhalten und Zielen von Denkprozessen.

Im realen sozialen Leben jedoch zählen sowohl Inhalt als auch Genauigkeit. Genauso wichtig wie das Verständnis der Prozesse, durch die sich Einstellungen, Eindrücke und Wissensstrukturen ausbilden, sie verschlüsselt, modifiziert, aktiviert und assoziiert werden, ist die Erforschung von tatsächlichen Inhalten dieser Strukturen. Beispielsweise mögen zwei verschiedene Eigenschaftskonzepte (z.B. maskulin, männlich) als Stimulus in einem Experiment austauschbar sein, da sie die gleiche Auswirkung auf den Prozess der Eindrucksbildung haben und vielleicht sogar eine sehr ähnliche evaluative Zuschreibung evozieren. Werden diese Eigenschaften allerdings in einen tatsächlichen Persönlichkeitseindruck kodiert, können sie einen ganz unterschiedlichen Einfluss auf die folgende Interaktion und damit auf die soziale Realität haben.

An diesem Punkt setzt die evolutionäre Sozialpsychologie an. Ihr spezifischer Ausgangspunkt ist, dass Denken (und Fühlen) einzig dem anschließenden Handeln dient. Eine spezifische Ausprägung dieser proximalen Mechanismen wiederum bleibt nur dann erhalten, wenn sie dem ultimativen Ziel, die Weitergabe der eigenen Gene, zuträglich ist. Mit anderen Worten sind es in allen sozialen Situationen wahrscheinlich die spezifisch fitnessrelevanten Informationen, denen die größte Aufmerksamkeit zuteil wird und die vorrangig verarbeitet werden, so dass das eigene Überleben und die eigene Fortpflanzung

wahrscheinlicher wird – der Inhalt zählt eben doch (Kenrick et al., 2006). Ein evolutionärer Ansatz hilft, soziale Situationen zu identifizieren, die besonders relevant für den Menschen sind und spezifische Formen von Kognitionen, Emotionen und Verhaltensweisen auslösen. Der evolutionäre Fokus liegt dabei auf mentalen Prozessen, die speziell angepasst sind, um die spezifischen Probleme relevanter sozialer Situationen zu lösen. Zu den direkt relevanten sozialen Problemen in Bezug auf Überleben und Reproduktion gehören Schutz des eigenen Lebens, Partnerwahl und Partnerbindung, Fürsorge unter Verwandten, Gruppenzugehörigkeit (Schaller et al., 2007) und relativer Status innerhalb der Gruppe.

Im Zuge dessen ist es wahrscheinlich, dass der Mensch – genauso wie er schnell gute von schlechten potenziellen Partnern unterscheiden kann – die Fähigkeit entwickelt hat, die potenziellen Auswirkungen anderer Personen auf den eigenen relativen sozialen Status, zuverlässig abzuschätzen. Zur Konfliktvermeidung und für die Erreichung der eigenen Ziele, sollte die Einschätzung von Führungsstärke ausreichend akkurat sein, da sie die Grundlage für angemessenes Folgeverhalten darstellt.

Der wesentliche Unterschied zwischen der klassischen kognitiven und der evolutionären Sozialpsychologie besteht somit in den gestellten Fragen. Während die evolutionäre Psychologie vor allem die Frage nach dem Ergebnis der sozialen Wahrnehmung stellt, beschäftigt sich die soziale Kognitionsforschung insbesondere mit dem Prozess der Eindrucksbildung.

Aus diesem Grund lässt sich aus der sozialen Kognitionsforschung für die Frage, ob Führungsstärke intuitiv erkennbar ist, keine gerichtete Annahme aufstellen. Aus der evolutionären Sozialpsychologie heraus wird die Hypothese formuliert, dass intuitive Beurteiler die Führungsstärke fremder Personen besser als zufällig einschätzen können.

Doch wie gelangen die Beurteiler zu ihrer Einschätzung?

Bezogen auf den Prozess der Zuschreibung von Führungsstärke können aus den unterschiedlichen Forschungsrichtungen drei einander widersprechende Hypothesen aufgestellt und gegeneinander getestet werden.

Aus der klassischen kognitiven Sozialpsychologie kann die Hypothese aufgestellt werden, dass sowohl das „think manager - think male“- (Powell et al., 2002; Schein, 1973; Sczesny et al., 2006) als auch das „what is beautiful is good“- Stereotyp (Dion et al., 1972; Feingold, 1992; Langlois et al., 2000) jeweils direkt die Zuschreibung von Führungsstärke beeinflussen, so dass Männern und attraktiveren Zielpersonen eine höhere Führungsstärke zugeschrieben wird.

Zweitens resultiert aus dem „Lack-of-fit“- Modell (Heilman, 1983; Heilman & Stopeck, 1985) die Hypothese, dass die beiden Zielpersonenmerkmale in einer bestimmten Interaktion ihren Einfluss ausüben. Hier lautet die Hypothese, dass attraktiven männlichen Zielpersonen eine hohe Führungsstärke zugesprochen wird, während attraktive weibliche Zielpersonen demgegenüber in der Wahrnehmung abfallen.

Auch aus evolutionärer Perspektive heraus wird angenommen, dass sich die Beurteiler an den Merkmalen Geschlecht und Attraktivität orientieren. Aus der Forschung zum „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007), kann drittens die Hypothese aufgestellt werden, dass eine Interaktion zwischen Geschlecht und Attraktivität auf Seiten der Zielperson abhängig von dem Geschlecht des Beurteilers auftritt. Bei der gegengeschlechtlichen Beurteilung müssten die attraktiven Zielpersonen besonders vorteilhaft bewertet werden, während gleichgeschlechtlich die mittelattraktiven die beste Bewertung erhalten sollten.

Würden schließlich die Beurteiler sowohl die Geschlechts- als auch die Attraktivitätsheuristik nutzen und würden sie auf diese Weise zu guten Einschätzungen gelangen, da diese ein „Körnchen Wahrheit“

besäßen, so könnte dies als Unterstützung für das Konzept der simplen und sparsamen Heuristiken gewertet werden.

6. Empirische Überprüfung

Um die formulierten Hypothesen mit ihren verschiedenen Facetten zu analysieren, wurden vier aufeinander aufbauende Studien durchgeführt. Allen Studien ist das Messinstrument für Führungsstärke, welches zu einem späteren Zeitpunkt näher erläutert wird, und das Untersuchungsparadigma der „Thin slices“- Forschung gemein. In allen Studien sahen intuitive Beurteiler einen Film, bestehend aus einer Anzahl von zuvor auf Video aufgenommenen Zielpersonen. Jede einzelne dieser Zielpersonen war 20 Sekunden lang ohne Ton beobachtbar und wurde von den Beurteilern auf ihre Führungsstärke oder auf ihre Attraktivität hin beurteilt. Zur Ermittlung der Genauigkeit der Personenwahrnehmung wurden jeweils die aggregierten Einschätzungen mit einem Kriterium verglichen, welches ebenso wie die Stichprobe der Beurteiler und die der Zielpersonen von Studie zu Studie variierte.

6.1. Überblick über die empirischen Studien

Das biologische Geschlecht und physische Attraktivität von Personen sind erstens saliente, schnell zugängliche und in vielen Kontexten relevante Anhaltspunkte für die Personenbewertung und zweitens Bestandteile von weit verbreiteten und gängigen Stereotypen. Daher wurde angenommen, dass beide Merkmale Auswirkungen auf die Zuschreibung von Führungsstärke haben. Über die Art und Weise dieses Einflusses wurden aus den verschiedenen relevanten Forschungsrichtungen konkurrierende Hypothesen aufgestellt.

Zum einen könnten direkte und lineare Einflüsse des Geschlechts- sowie des Attraktivitätsstereotyps auftreten, so dass Männern und attraktiven Zielpersonen eine höhere Führungsstärke zugeschrieben

wird. Zweitens könnte sich eine Interaktion zwischen den Merkmalen Geschlecht und Attraktivität auf Seiten der Zielpersonen aufzeigen lassen, indem attraktiven männlichen Zielpersonen eine hohe, attraktiven weiblichen Zielpersonen dagegen eine niedrigere Führungsstärke attestiert wird („Lack-of-fit“-Modell). Drittens wäre eine Interaktion zwischen Geschlecht und Attraktivität auf Seiten der Zielperson in Abhängigkeit von dem jeweiligen Geschlecht des Beurteilers möglich. Im Falle eines „sexual attribution bias“ müsste bei der gegengeschlechtlichen Einschätzung attraktive Zielpersonen die höchsten (Glorifizierungskomponente), bei der gleichgeschlechtlichen jedoch die mittelattraktiven Zielpersonen die höchsten Bewertungen erhalten (Abwertungskomponente).

Ob die beiden Merkmale Geschlecht und Attraktivität bei der Zuschreibung von Führungsstärke genutzt werden und welches dieser drei Phänomene bei der Zuschreibung von Führungsstärke auftritt, wurde mit Hilfe von Studie 1 untersucht.

Über die stereotype Wahrnehmung von Führungsstärke hinaus wurde in Studie 2 der Frage nachgegangen, ob Menschen in der Lage sind, die tatsächliche Führungsstärke anderer, fremder Personen besser als zufällig zu erkennen. Aus dem bisherigen Forschungsstand ableitbare Vermutungen zu dieser Frage stellen sich keinesfalls als eindeutig dar. Da angenommen wird, dass die Beurteiler im Zuschreibungsprozess auf die Merkmale Geschlecht und Attraktivität zurückgreifen, könnte man argumentieren, dass eine akkurate Wahrnehmung nicht möglich sei, wenn Beurteiler ihre Einschätzung auf Stereotype basieren. Andererseits resultiert aus dem evolutionspsychologischen Ansatz und den Ergebnissen der ABC Forschergruppe die Vermutung, dass intuitive Urteile auf Basis simpler und sparsamer heuristischer Strategien getroffen werden, die dennoch zum gewünschten Ziel führen. Studie 2 analysierte somit die Frage, ob Genauigkeit in der Personenwahrnehmung trotz oder gerade we-

gen der Anwendung von Stereotypen erreicht werden kann und ob diese Stereotype ein „Körnchen Wahrheit“ besitzen.

Die Durchführung von Studie 3 diente der Erforschung eines weiteren zentralen Aspektes der allgemeinen Fragestellung. Hier wurden die Einschätzungen von studentischen (Laien-) Einschätzungen über Führungsstärke mit dem Urteil von Personalverantwortlichen verglichen, um so zu untersuchen, ob Personenwahrnehmung trainiert werden kann und inwiefern Zuschreibungsprozesse sich im Laufe des Lebens und mit mehr Erfahrung verändern. Studie 3 diente somit dem vorrangigen Ziel, die Rolle von Expertise im intuitiven Beurteilungsprozess zu untersuchen.

Studie 4 diente insbesondere der Erweiterung des Geltungsbereiches der Ergebnisse. Hier wurde auf Seiten der Zielpersonen über eine studentische Stichprobe hinausgegangen und eine Stichprobe von Arbeitnehmern rekrutiert. Diese Maßnahme erlaubte gleichsam andere Referenzkriterien für Führungsstärke, nämlich ein 360°-Feedback, sowie, als verdichtetes Maß für Führungsstärke, die tatsächliche berufliche Position. Studie 4 erlaubte somit die Übertragung der Ergebnisse auf eine Stichprobe mit unterschiedlicher Altersstruktur, beruflichem Hintergrund und auf eine andere Art und Weise gemessene Führungsstärke.

Führungsstärke wurde im Sinne von charismatischer Führung als ein soziales Charakteristikum definiert, welches einerseits auf einer Führungsdisposition und gleichzeitig auf konkreten führungsrelevanten Verhaltensweisen beruht. Ihre Messung stellt daher einen diffizilen und dennoch entscheidenden Faktor für die Aussagekraft der Ergebnisse dar.

6.2. Zur Messung von Führungsstärke

Als Operationalisierung für den dispositionellen Anteil von Führungsstärke diente in allen Studien die vom Allensbacher Institut unter der

Leitung von Elisabeth Noelle-Neumann (1983) entwickelte Skala Persönlichkeitsstärke.

Die Ursprungsversion dieses Fragebogens besteht aus insgesamt 13 Items, davon 10 Selbsteinschätzungsfragen und 3 Fragen zu objektiven Daten, welche allesamt mit „trifft zu“ oder „trifft nicht zu“ zu beantworten sind (Anhang A).

Die Forschung Noelle-Neumanns unter Verwendung dieser Skala zeigte, dass diese in der Lage ist, Führungsstärke als Persönlichkeitsdisposition unabhängig von Alter und sozioökonomischem Status zu erfassen. Persönlichkeitsstärke bezeichnet laut Noelle-Neumann eine besondere Ausstrahlung („Mana“), „eine psychische Kraft, die manche Menschen besitzen; sie ist in ihnen, gehört zum Charakter, kann nicht von außen beeinflusst oder verändert werden.“ (Noelle-Neumann, 1983, S. 7). Mit dem Begriff „Mana“ bezeichnete sie damals eine mitreißende, natürliche und besondere Ausstrahlung. Diese findet man auch heute in der modernen Führungsforschung, hier trägt sie allerdings den Namen Charisma.

Noelle-Neumann wies anhand einer repräsentativen Bevölkerungsstichprobe nach, dass Menschen mit einer hohen Persönlichkeitsstärke sich in vielen Aspekten grundsätzlich von Menschen mit niedriger Persönlichkeitsstärke unterscheiden. So haben sie einen größeren Bekanntenkreis, werden öfter von ihren Bekannten, Kollegen und Freunden um Rat gefragt, üben mehr Einfluss auf andere aus, sind offener für Neues, zeigen eine Bereitschaft, sich ständig weiterzuentwickeln, üben lehrende Tätigkeiten in ihrer Freizeit aus (geben Kurse, trainieren Mannschaften, etc.), sind selbstkritisch, weisen eine optimistische und lebensfrohe Grundeinstellung auf, empfinden eine große Freude an der Arbeit und engagieren sich überdurchschnittlich in Beruf und Freizeit (Noelle-Neumann, 1983).

Zusammenfassend bedeutet dies, dass persönlichkeitsstarke Menschen in privaten wie auch beruflichen Situationen eher nach Führungsrollen streben als persönlichkeitschwache Personen. Das

aktive Aufsuchen von Verantwortung ist nach Weiss (2003) ein Merkmal wahrer Führungsstärke. Gleichzeitig zeigen diese Ergebnisse, dass Menschen sich von persönlichkeitsstarken Personen im privaten wie auch beruflichen Bereich gerne anleiten, beeinflussen und führen lassen. Auch bei charismatischer Führung ist die Identifikation mit dem Führenden und das Ausrichten auf dessen Ziele ein zentraler Bestandteil (Yukl, 2002).

Somit eignet sich die Skala als Kriterium zur Untersuchung von Führungsstärke und dies insbesondere auch bei Studierenden, die eine eigenschaftsbasierte Disposition für Führungsstärke aufweisen können, welche sich jedoch noch nicht in einem entsprechenden beruflichen Status äußert.

Dass es in der Tat einen Zusammenhang zwischen Persönlichkeitsstärke und beruflicher Position gibt, fand Noelle-Neumann ebenfalls in ihrer Studie von 1983. Arbeitnehmer mit einer hohen Persönlichkeitsstärke bekleideten mit einer höheren Wahrscheinlichkeit eine berufliche Führungsposition als Personen mit einer niedrigen Persönlichkeitsstärke. Von den untersuchten Personen aus den oberen Führungsebenen wiesen etwa ein Viertel eine hohe bis sehr hohe Persönlichkeitsstärke auf, keiner dagegen eine sehr niedrige. Auf der anderen Seite hatten aus der Gruppe, die auf der Skala sehr niedrige Werte erreichten, 99% keine berufliche Führungsposition und lediglich 1% eine mittlere Führungsposition. Diese Befunde können als ersten Hinweis für die externe Validität der Skala gewertet werden.

Ein weiterer Hinweis stammt aus einer Studie von Weimann (1991). Er fand einen Zusammenhang zwischen Persönlichkeitsstärke und der Position im Kommunikationsnetz einer israelischen Kibbuz-Gemeinde. Zunächst erstellte er eine „soziometrische Karte“, die der Darstellung des persönlichen Kommunikationsnetzes, des Informationsflusses und des persönlichen Einflusses innerhalb der Gemeinde diente. Anhand dieser Karte verglich er die Zentralität einer Person in der Gemeinde, gemessen durch die Anzahl ihrer Verbindungspunkte

mit anderen Mitgliedern, mit ihren Persönlichkeitsstärkewerten. Als Ergebnis zeigte sich ein statistisch signifikanter Zusammenhang von $r = .54$ zwischen Persönlichkeitsstärke und der Anzahl an Verbindungspunkten im Netzwerk. Persönlichkeitsstarke Gemeindemitglieder hatten nicht nur eine zentralere Position im Kommunikationsnetz inne, sie waren auch aktiver am Informationsfluss (mit Ausnahme von Klatsch) und an Entscheidungsfindungen beteiligt (Weimann, 1991).

Diese Ergebnisse zeigen, dass die Skala Persönlichkeitsstärke ein Instrument darstellt, um Personen zu identifizieren, die - unabhängig von Schicht, Alter und sozioökonomischem Hintergrund - eine höhere gesellschaftliche Position innehaben, einen meinungsführenden Einfluss auf andere ausüben und die sowohl beruflich als auch privat andere Menschen überdurchschnittlich oft führen.

Konkret wurden für die vorliegenden Studien drei Items der Skala ausgewählt und in die dritte Person umformuliert. Es handelte sich um die Fragen (1) „Person rechnet bei dem was sie macht, gewöhnlich mit Erfolg.“, (2) „Person übernimmt bei gemeinsamen Unternehmungen gerne die Führung.“ und (3) „Person kann sich gut durchsetzen.“ Dies geschah, da die Beurteiler zwischen den einzelnen Videosequenzen nicht genug Zeit hatten, um alle 13 Fragen zu beantworten.

Die Auswahl der Items erfolgte nach drei Gesichtspunkten. Erstens gehören diese drei Items zu den von Weimann (1991) als „internale Faktoren“ der Einflussnahme auf andere bezeichneten Dimension, welche abzugrenzen sind von den „externalen“, hauptsächlich auf dem Vergleich mit anderen Personen beruhenden Quellen von Einflussnahme. Zweitens zeichnen sich diese drei Fragen durch eine hohe Augenscheinvalidität aus, das heißt, sie sind für die Beurteiler intuitiv besonders zugänglich. Drittens stellen diese drei Fragen den Kern des zu messenden Konstrukts Führungsstärke dar, indem sie die Dimensionen Selbstsicherheit, Durchsetzungsfähigkeit und die

allgemeine Bereitschaft, Führungsrollen zu übernehmen (Weiss, 2003), erfassen.

Statt der Antwortmöglichkeiten „Trifft zu“ beziehungsweise „Trifft nicht zu“ wurde den Beurteilern eine siebenstufige Antwortskala vorgegeben, um die Varianz der Antworten zu erhöhen und eine größtmögliche Differenzierung der Zielpersonen zu ermöglichen.

In allen folgenden Studien beruht die von den Intuitivbeurteilern zugeschriebene Führungsstärke jeder Zielperson somit auf den gemittelten Antworten zu diesen drei Items.

6.3. Studie 1

Es wurde vermutet, dass die Einschätzung der intuitiven Beurteiler sowohl vom Geschlechts- als auch vom Attraktivitätsstereotyp beeinflusst wird. Studie 1 diene der Frage, ob die Vorhersagen der klassischen kognitiven Sozialpsychologie, des „Lack-of-fit“- Modells oder des „sexual attribution bias“ zur Richtung des Einflusses von Geschlecht und Attraktivität bei der intuitiven Wahrnehmung von Führungsstärke auftreten.

Die kognitive Sozialpsychologie würde einen direkten und linearen Einfluss beider Merkmale auf die zugeschriebene Führungsstärke voraussagen. Männern und attraktiven Zielpersonen würden entsprechend des „think manager – think male“- Phänomens (Schein, 1973; Sczesny, 2003b; Sczesny et al., 2006) beziehungsweise der „what is beautiful is good“- Heuristik (Dion et al., 1972; Feingold, 1992; Langlois et al., 2000) eine höhere Führungsstärke zugeschrieben werden.

Das „Lack-of-fit“- Modell (Heilman, 1983; Heilman & Stopeck, 1985) dagegen würde auf Seiten der Zielpersonen eine Interaktion zwischen Geschlecht und Attraktivität voraussagen. In diesem Falle würde bei attraktiven männlichen Zielpersonen eine hohe, bei attraktiven weib-

lichen Zielpersonen dagegen eine niedrigere Führungsstärke wahrgenommen.

Schließlich ginge der „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007) ebenfalls von einer Interaktion zwischen Geschlecht und Attraktivität aus, wobei das spezielle Zuschreibungsmuster von dem Geschlecht des Beurteilers abhinge. In diesem Falle müsste bei der gegengeschlechtlichen Einschätzung ein linearer Zusammenhang zwischen Attraktivität und zugeschriebener Führungsstärke zu finden sein, so dass die Zielpersonen mit steigender Attraktivität höhere Bewertungen erhalten (Glorifizierungskomponente). Bei der gleichgeschlechtlichen Bewertung jedoch sollte es zu einem kurvilinearen Zusammenhang zwischen Attraktivität und zugeschriebener Führungsstärke kommen, so dass bei mittelattraktiven Zielpersonen die höchste Führungsstärke wahrgenommen würde (Abwertungskomponente).

6.3.1.Methode

Zielpersonen

In einem ersten Schritt wurde das Stimulusmaterial für die Studie generiert. Zu diesem Zweck wurden auf dem Campus der Universität Groningen 151 Studierende angeworben, davon 45 der Fachrichtung Biologie, 50 Psychologie- und 56 Betriebswirtschaftsstudenten. Im Versuchsraum wurden alle Teilnehmer zunächst gebeten, sich kurz in die laufende Kamera vorzustellen. Die Dauer der Aufnahme betrug etwa eine Minute. Alle Zielpersonen saßen dabei an einem Tisch und blickten geradeaus in die Kamera, so dass jeweils der gleiche Bildausschnitt, bestehend aus Kopf, Oberkörper und Armen, entstand.

Die generierten Videoclips wurden anschließend pro Zielperson jeweils auf 20 Sekunden gekürzt und in einen tonlosen Gesamtfilm verarbeitet, wobei einer Sequenz von 20 Sekunden jeweils eine 9-sekündige Pause folgte. Während dieser Pause war der Bildschirm schwarz, der nächste Filmausschnitt wurde mit einem Ton angekün-

digt. Die Gesamtdauer des Films mit allen 151 Zielpersonen betrug 75 Minuten.

Die Zielpersonen wurden auf Basis dieser Videosequenzen von 5 weiblichen und 7 männlichen Studenten der Universität zu Köln im Alter von 19 bis 37 Jahren ($M = 27.75$, $SD = 6.27$) auf ihre Attraktivität hin beurteilt. Die Frage für die Beurteiler lautete „Person 1-151 finde ich...“ und wurde auf einer siebenstufigen Skala beantwortet. Die durchschnittliche Attraktivität lag bei 3.54 ($SD = 1.09$). Ein Intraklassen-Koeffizient (ICC) von .42 führte zu einem Cronbach's Alpha von .90, so dass die einzelnen Attraktivitätsurteile zu einem Wert pro Zielperson gemittelt werden konnten.

Beurteiler

Insgesamt 42 Studenten der Psychologie und der Betriebswirtschaftslehre der Universität zu Köln nahmen als Beurteiler an Studie 1 teil, davon 22 weiblich und 20 männlich. Die Beurteiler befanden sich im Alter von 18 bis 37 Jahren, im Durchschnitt 25.15 Jahre ($SD = 5.11$).

Versuchsaufbau

Die intuitiven Beurteiler sahen den tonlosen Gesamtfilm der gesamten 151 Zielpersonen. Vorab wurden sie vom Versuchsleiter in das Ziel und den Ablauf der Untersuchung eingeführt. Ihnen wurde mitgeteilt, dass es sich um eine Studie zum ersten Eindruck handele und sie wurden ausdrücklich darauf hingewiesen, dass sich das Interesse der Untersuchung ausschließlich darauf richte, wie sie die folgenden Personen nach ihrem ersten Eindruck einschätzen. Anschließend wurden die Fragebögen ausgeteilt. Die schriftlichen Instruktionen betonten erneut, dass es sich um eine Studie zum ersten Eindruck handele und dass es keine richtigen oder falschen Antworten gebe.

Die Aufgabe der Probanden bestand darin, nach jedem einzelnen Videoclip die soeben gesehene Zielperson anhand der ihnen vorliegenden Fragen zu beurteilen. Die drei Fragen zur Einschätzung der

Führungsstärke wurden dabei der Skala Persönlichkeitsstärke nach Noelle-Neumann (1983) entnommen und lauteten: (1) „Person 1-151 rechnet bei dem was sie macht, gewöhnlich mit Erfolg.“, (2) „Person 1-151 übernimmt bei gemeinsamen Unternehmungen gerne die Führung.“ und (3) „Person 1-151 kann sich gut durchsetzen.“ Alle Fragen wurden jeweils auf einer siebenstufigen Antwortskala erhoben (Anhang B).

Nach der Hälfte des Gesamtfilmes wurde eine Pause von zehn Minuten eingelegt. Abschließend wurde den Teilnehmern gedankt und sie wurden über das Ziel der Studie aufgeklärt.

6.3.2. Ergebnisse

Bei der Analyse der Daten wurde zunächst der Frage nachgegangen, inwieweit sich die Beurteiler in ihrer Einschätzung der Persönlichkeitsstärke der Zielpersonen einig waren. Eine Reliabilitätsanalyse ergab einen Intraklassen-Koeffizient (*ICC*) von .24, welcher zu einem Cronbach's Alpha von .93 führte. Die Interraterreliabilität war somit als ausreichend zu bewerten und erlaubte die Aggregation der Einzelwerte zu einem Gesamtmittelwert. Außerdem waren alle Korrelationen der Intuitivurteile untereinander positiv.

Weiterhin lässt sich festhalten, dass die Zielpersonen systematisch unterschiedlich eingeschätzt wurden. Bei zufälligen Schätzungen sollte allen Zielpersonen in etwa eine gleich hohe Persönlichkeitsstärke zugeschrieben werden. Dies war jedoch nicht der Fall. Die Zielperson, die von den Beurteilern als am wenigsten persönlichkeitsstark wahrgenommen wurde, erreichte eine durchschnittlich wahrgenommene Persönlichkeitsstärke von 2.55. Die Zielperson dagegen, der die höchste Persönlichkeitsstärke attestiert wurde, erreichte einen durchschnittlichen Schätzwert von 5.50.

Nachdem die Einigkeit der Beurteiler und die systematische Variation der Antworten als Grundlage für die weitere Analyse festgestellt wurden, konnte in einem nächsten Analyseschritt der Frage nachgegan-

gen werden, ob und wenn ja wie das Attraktivitäts- und das Geschlechtsstereotyp die Personenwahrnehmung beeinflusst.

Auf bivariater Ebene zeigte sich, dass die Attraktivität mit den Einschätzungen von Persönlichkeitsstärke signifikant korrelierte ($r = .51$, $p < .001$). Die Attraktivität der Zielpersonen wurde somit bei der Urteilsbildung berücksichtigt, insofern als dass attraktiveren Personen eine höhere Persönlichkeitsstärke attestiert wurde.

Hinsichtlich des Geschlechtes zeigte sich jedoch kein signifikanter Zusammenhang mit der zugeschriebenen Persönlichkeitsstärke ($r = -.09$, n.s.).

Allerdings zeigte sich ebenfalls, dass Attraktivität der Zielpersonen nicht nur signifikant mit der wahrgenommenen Persönlichkeitsstärke korrelierte, sondern ebenfalls mit der Variable Geschlecht mit $r = .34$ ($p < .001$). Wenn Attraktivität in einem Zusammenhang mit der wahrgenommenen Führungsstärke steht und Frauen attraktiver wirken als Männer, dann sollten unter Konstanthaltung des Einflusses der Attraktivität weibliche Zielpersonen bezogen auf die ihnen zugeschriebene Persönlichkeitsstärke gegenüber Männern abfallen.

Eine Betrachtung der Zusammenhänge auf multivariater Ebene enthüllte diesen vermuteten Suppressor-Effekt. Es wurde eine Regressionsanalyse mit der abhängigen Variablen zugeschriebene Persönlichkeitsstärke und den unabhängigen Variablen Geschlecht und Attraktivität berechnet.

Hier zeigte sich im ersten Schritt, dass beide Prädiktoren unabhängig voneinander mit einem β von $.61$ für Attraktivität und von $-.29$ für Geschlecht einen signifikanten Einfluss ($p < .001$) auf die zugeschriebene Persönlichkeitsstärke hatten. Männlichen und attraktiveren Zielpersonen wurde eine höhere Persönlichkeitsstärke zugeschrieben. Beide Merkmale zusammen konnten etwa 34% der Varianz des Antwortverhaltens der Beurteiler erklären.

Im zweiten Schritt der Regression wurde zusätzlich zu den beiden Prädiktoren Geschlecht und Attraktivität der Interaktionsterm aus Geschlecht und Attraktivität hinzugenommen, um die aus dem „Lack-of-fit“-Modell heraus aufgestellte Hypothese eines Zusammenwirkens beider Merkmale auf Seiten der Zielpersonen zu prüfen. Der Interaktionsterm wurde zuvor durch Zentrierung und anschließender Multiplikation der beiden Variablen Attraktivität und Geschlecht gebildet. Bei diesem zweiten Analyseschritt war jedoch durch die Hinzunahme der Interaktion der beiden Prädiktorvariablen kein Zuwachs in der aufgeklärten Varianz zu verzeichnen (Änderung in $F = .00$, Änderung in $R^2 = .00$, n.s.). Diese blieb konstant bei 34%. Somit konnte eine Interaktion der beiden Merkmale Attraktivität und Geschlecht auf Zielpersonenseite nicht bestätigt werden.

Im Folgenden wurde überprüft, ob sich die dritte Hypothese als zutreffend erwies, das heißt ob ein „sexual attribution bias“ vorlag. Zu diesem Zweck wurden die Daten auf ein Zusammenwirken der Merkmale Geschlecht und Attraktivität in Abhängigkeit von dem Geschlecht der Zielperson und dem Geschlecht der Beurteiler analysiert. Dazu wurden jeweils Korrelationen zwischen der Attraktivität der Zielpersonen und der ihnen zugeschriebenen Persönlichkeitsstärke, getrennt nach dem Geschlecht der Beurteiler und dem Geschlecht der Zielperson, berechnet.

Bei männlichen Zielpersonen zeigte sich, dass der Zusammenhang zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke signifikant positiv ausfiel, egal ob das Urteil von weiblichen Beurteilern ($r = .54$, $p < .001$) oder von männlichen Beurteilern ($r = .52$, $p < .001$) stammte. Das gleiche Bild zeigte sich bei weiblichen Zielpersonen. Männliche und weibliche Intuitivbeurteiler schrieben attraktiveren weiblichen Zielpersonen eine höhere Persönlichkeitsstärke zu. Dieser Zusammenhang lag bei männlichen Beurteilern bei $r = .60$ ($p < .001$) und bei weiblichen bei $r = .55$ ($p < .001$). Weder die Zusammenhänge in Bezug auf die Einschätzung der männlichen ($Z = 0.12$, n.s.),

noch auf die weiblichen Zielpersonen ($Z = -0.33$, n.s.) unterschieden sich signifikant.

Obwohl diese Ergebnisse bereits einen Hinweis auf die Adäquatheit der Annahme eines linearen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke geben, wurde als letzte Maßnahme zur Überprüfung des „sexual attribution bias“ Kurvenanpassungen mit der unabhängigen Variablen Attraktivität der Zielpersonen und der abhängigen Variablen zugeschriebene Persönlichkeitsstärke berechnet. Auch diese Berechnungen erfolgten jeweils getrennt nach Geschlecht der Beurteiler und der Zielpersonen. So resultierten insgesamt die vier Kombinationen (1) Einschätzung der weiblichen Zielpersonen durch weibliche Beurteiler, (2) durch männliche Beurteiler, (3) Einschätzung der männlichen Zielpersonen durch weibliche Beurteiler und (4) durch männliche Beurteiler.

Prinzipiell könnte ein vorhandener kurvilinearere Zusammenhang durch die spezielle Zusammensetzung der Stichprobe verschleiert werden, wenn das verwendete Stimulusmaterial nicht das gesamte Attraktivitätsspektrum abbildet. Sollten die als Stimulusmaterial dienenden Zielpersonen alle ungefähr durchschnittlich attraktiv sein, so könnte es sein, dass die Annahme eines linearen Zusammenhangs nur durch das Fehlen der Extrempunkte sehr unattraktiv und sehr attraktiv adäquat erscheint. Zur Illustration zeigen die Abbildungen 1-4, dass in allen vier Kombinationen die Attraktivität der Zielpersonen das gesamte Spektrum abbildete.

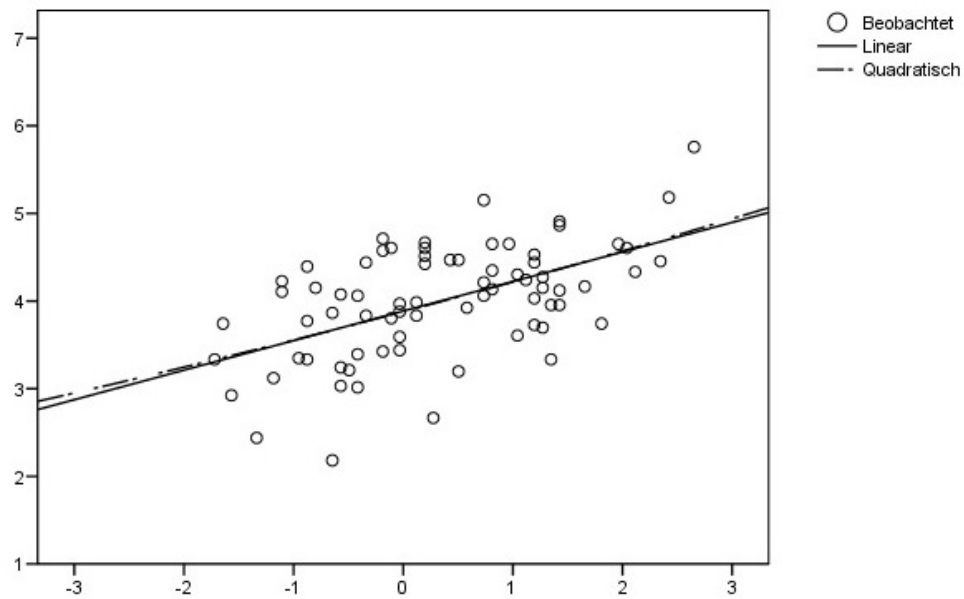


Abbildung 1. Streuung der Urteile über Persönlichkeitsstärke durch weibliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der weiblichen Zielpersonen.

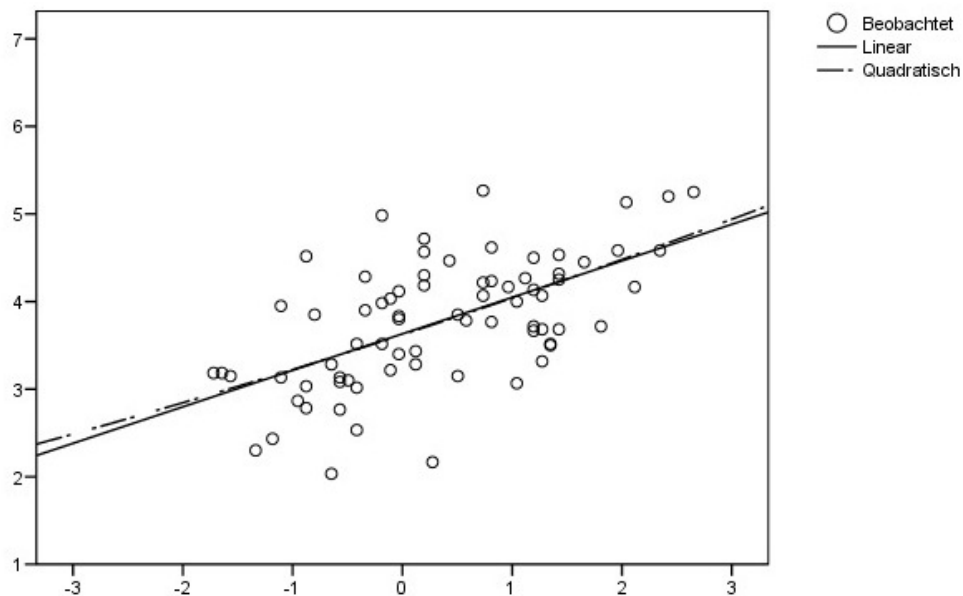


Abbildung 2. Streuung der Urteile über Persönlichkeitsstärke durch männliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der weiblichen Zielpersonen.

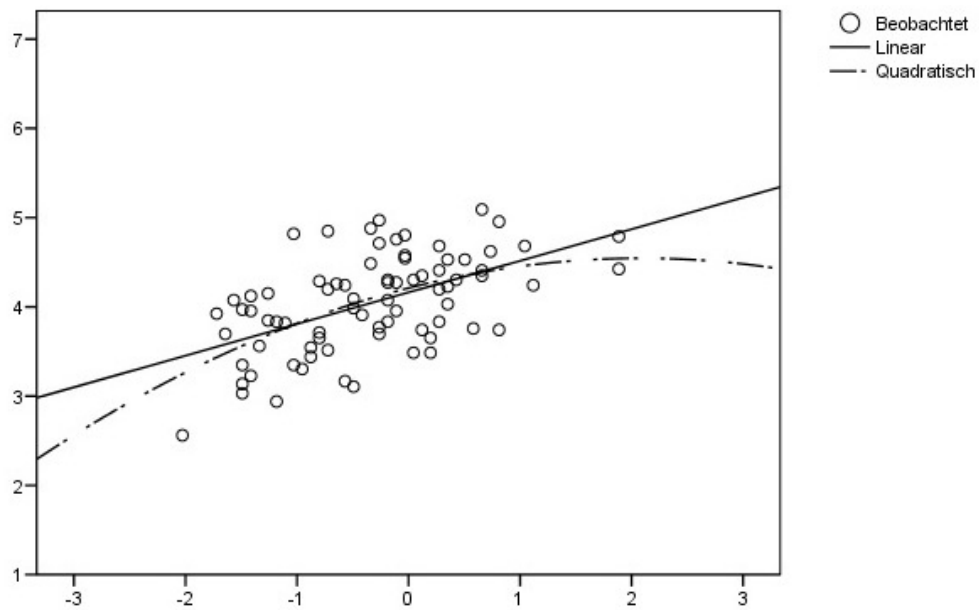


Abbildung 3. Streuung der Urteile über Persönlichkeitsstärke durch weibliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der männlichen Zielpersonen.

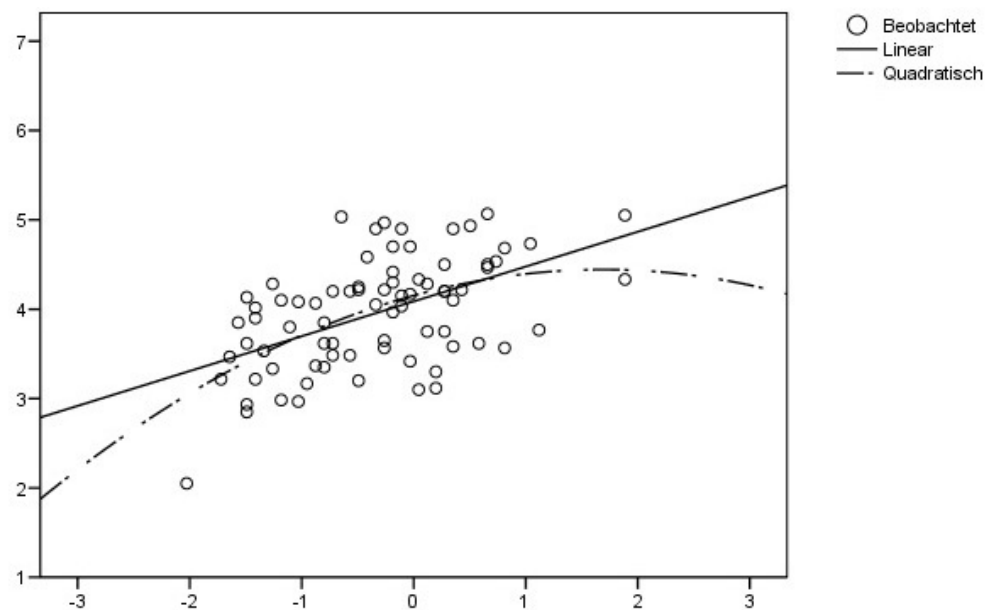


Abbildung 4. Streuung der Urteile über Persönlichkeitsstärke durch männliche Beurteiler in Abhängigkeit von der zentrierten Attraktivität der männlichen Zielpersonen.

Tabelle 3 zeigt die aufgeklärte Varianz in allen vier Kombinationen jeweils bei Annahme eines linearen im Vergleich zu einem kurvilinearen Zusammenhang. Ein Steigers Z-Test (Steiger & Brown, 1984; Streukens & De Ruyter, 2004) wies in keinem der Fälle eine signifi-

kant bessere Modellanpassung des quadratischen gegenüber dem linearen Modell aus. Läge ein „sexual attribution bias“ vor, so müsste jeweils bei der gleichgeschlechtlichen Beurteilung die aufgeklärte Varianz im kurvilinearen Modell deutlich ansteigen. Da dies nicht gezeigt werden konnte, lag kein „sexual attribution bias“ vor.

Zusammenhang	<i>weibliche Beurteiler, weibliche Zielpersonen</i>	<i>Weibliche Beurteiler, männliche Zielpersonen</i>	<i>Männliche Beurteiler, weibliche Zielpersonen</i>	<i>Männliche Beurteiler, männliche Zielpersonen</i>
Linear	30%	29%	36%	27%
Quadratisch	30%	31%	36%	30%
Steigers Z	.00, n.s.	-.73, n.s.	.00, n.s.	-1.43, n.s.

Tabelle 1. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielperson (*: $p < .05$; **: $p < .01$).

6.3.3.Diskussion

Mit Hilfe dieser ersten Studie wurde den Fragen nachgegangen, ob und in welcher Weise gängige Stereotype die intuitive Wahrnehmung von Führungsstärke beeinflussen.

Ein erstes festzuhaltendes Ergebnis ist, dass die Beurteiler sich sowohl bei der Beurteilung von Führungsstärke als auch von Attraktivität sehr einig zeigten. Diese Einigkeit konnte hinsichtlich Attraktivität in der Forschung bereits demonstriert werden (Langlois et al., 2000). Die intuitiven Beurteiler schrieben jedoch auch den gleichen Personen eine hohe Führungsstärke und den gleichen Personen eine niedrige Führungsstärke zu. Somit ist der Schluss erlaubt, dass die Beurteiler sich an Charakteristiken der Zielpersonen orientieren.

Über die Art und Weise, wie das Geschlechts- und das Attraktivitätsstereotyp die Zuschreibung von Führungsstärke beeinflussen könnten, wurden vorab drei konkurrierende Hypothesen aufgestellt.

Bestätigt werden konnte zunächst die Annahme, dass sowohl die Attraktivität der Zielpersonen gemäß der „what is beautiful is good“-Heuristik (Dion et al., 1972; Feingold, 1992; Langlois et al., 2000) als auch das Geschlecht der Zielpersonen gemäß des „think manager - think male“- Stereotyps (Schein, 1973, 1975; Sczesny et al., 2006) jeweils direkte Auswirkungen auf den Wahrnehmungsprozess hatten. Männlichen sowie attraktiveren Zielpersonen wurde eine höhere Führungsstärke zugeschrieben. Dieser Effekt zeigte sich aufgrund eines Suppressor-Effektes von Attraktivität jedoch erst auf multivariater Ebene. Auf bivariater Ebene konnten weibliche Zielpersonen ihre als geringer wahrgenommene Führungsstärke zunächst durch ihre höhere Attraktivität ausgleichen.

Die beiden anderen Hypothesen konnten dagegen nicht bestätigt werden. Erstens veränderte sich die aufgeklärte Varianz des Antwortverhaltens der Beurteiler unter Einbezug der Interaktion zwischen Geschlecht und Attraktivität auf Zielpersonenseite nicht. Somit konnten die Voraussagen des „Lack-of-fit“- Modells (Heilman, 1983; Heilman & Stopeck, 1985) nicht bestätigt werden. Bei attraktiven weiblichen Personen wurde somit im Vergleich zu weniger attraktiven weiblichen Zielpersonen nicht weniger Führungsstärke wahrgenommen. Ebenso nahm die zugeschriebene Führungsstärke bei männlichen Zielpersonen mit steigender Attraktivität zu.

Zweites konnten auch die Vorhersagen des „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007) in den Daten nicht gefunden werden. Sowohl bei männlichen als auch bei weiblichen Beurteilern hing die zugeschriebene Führungsstärke von der Attraktivität der Zielpersonen ab. Die Urteile unterschieden sich dabei jedoch nicht in Abhängigkeit von dem Geschlecht der Zielpersonen, wie es von dem „sexual attribution bias“ vorhergesagt wurde.

So konnte die Annahme eines kurvilinearen Zusammenhangs zwischen Attraktivität und zugeschriebener Führungsstärke bei gleichgeschlechtlicher Beurteilung nicht aufgezeigt werden. Beurteilten Frauen Frauen und Männer Männer, so erwies sich die Annahme eines linearen Zusammenhangs zwischen der Attraktivität der Zielpersonen und der wahrgenommenen Führungsstärke als adäquat. Waren die Zielpersonen attraktiver, so wurde ihnen durch die Beurteiler mehr Führungsstärke attestiert, unabhängig davon, ob sie durch weibliche oder männliche Personen eingeschätzt wurden.

Über die gefundenen direkten Auswirkungen des Geschlechts- sowie des Attraktivitätsstereotyps hinaus kann anhand dieser Studie keine Aussage über die Genauigkeit der Wahrnehmung von Führungsstärke getroffen werden. Diese Frage wurde mit Studie 2 aufgegriffen.

6.4. Studie 2

Die Durchführung von Studie 2 diente zwei unterschiedlichen Zielen. Einerseits wurde schwerpunktmäßig der Frage nachgegangen, ob intuitive Beurteiler die Führungsstärke fremder Personen auf Basis von 20-sekündigen tonlosen Videosequenzen zutreffend einschätzen können. In diesem Zuge sollte auch überprüft werden, ob die Beurteiler über die Anwendung von Stereotypen hinaus auch die Führungsstärke der Zielpersonen erkennen konnten beziehungsweise ob in den Stereotypen eventuell ein „Körnchen Wahrheit“ liegen könnte. Darüber hinaus sollten die in Studie 1 getesteten Annahmen über die Art und Weise der Nutzung von Geschlechts- und Attraktivitätsstereotypen repliziert werden. Auch in Studie 2 wurden die drei deduktiv abgeleiteten, konkurrierenden Hypothesen aus der kognitiven Sozialpsychologie, des „Lack-of-fit“- Modells sowie des „sexual attribution bias“ auf ihre Vorhersagekraft hin geprüft.

6.4.1. Ein Referenzmaß für Führungsstärke: die Verhaltensbeobachtung

Eine große Schwierigkeit bei der Erforschung der Genauigkeit von intuitiven Eindrücken ist die Wahl eines passenden Kriteriums. Ohne einen angemessenen Referenzwert ist eine Aussage, ob eine Persönlichkeitseinschätzung richtig oder falsch ist, schlicht nicht möglich. Doch welches Kriterium stellt im vorliegenden Fall das sinnvollsten dar?

Häufig verwendete Ansätze im „Thin slices“-Paradigma sind die Referenzkriterien Konsens, Übereinstimmung und Vorhersage.

Die Annäherung an Genauigkeit über *Konsens* ist eine gängige Methode, die einer konstruktivistischen Weltsicht entspricht (Kenny, 1991). Das Maß für Genauigkeit ist in diesem Falle die Einigkeit der Beurteiler untereinander. Anhand dieser Sichtweise kann kein Urteil eines Beurteilers über eine Zielperson als „wahrer“, „richtiger“ oder „genauer“ angesehen werden als die Wahrnehmung eines anderen Beurteilers, da alle Interpretation im Grunde nur soziale Konstruktionen darstellen (Gergen, 2001). Nach dieser Sichtweise wird ein Individuum nur in seiner Zugehörigkeit zu einem sozialen Kontext real, da menschliches Empfinden und die Art und Weise wie Menschen wahrgenommen werden ausschließlich im sozialen Kontext verankert ist (Jacob, 2004). Die einzige Annäherung an Wahrheit kann unter diesen Umständen durch die kollektive Sichtweise einer Gemeinschaft von Beurteilern erfolgen. In Studie 1 wurde gezeigt, dass unter den intuitiven Beurteilern Konsens vorlag. Konsens als Referenzkriterium wird in der vorliegenden Studie aus folgendem Grund als nicht ausreichend erachtet.

Die Messung von Genauigkeit alleinig über den Konsens der Beurteiler negiert die Möglichkeit einer Realität außerhalb und unabhängig von der unmittelbaren Wahrnehmung der Beurteiler. Theoretisch könnten sich alle Beobachter beispielsweise lediglich an den gleichen Stereotypen und Vorurteilen orientieren, auf diese Art und Weise zu einheitlichen Urteilen gelangen und dadurch gegebenenfalls wieder-

um soziale Realität schaffen. So wurde in Studie 1 auch deutlich, dass die Beurteiler sich an gängigen Stereotypen wie dem Geschlechts- und Attraktivitätsstereotyp ausrichten. Nimmt man eine psychologische Realität an, die außerhalb der Subjektivität der Beurteiler existiert, dann wären gerade die Eigenschaften, die eine Person *wirklich* besitzt, diejenigen Merkmale, die zukünftiges Verhalten determinieren und somit einen gewissen Grad an Vorhersagbarkeit ermöglichen. Da eine fehlerbehaftete kollektive Wahrnehmung durch die Konsens-Methode nicht final ausgeschlossen werden kann, wurde von diesem Kriterium für Genauigkeit abgesehen.

Stattdessen wurde in Studie 2 und 3 das Kriterium Vorhersage herangezogen, in Studie 4 das Kriterium Übereinstimmung.

Als *Vorhersage* bezeichnet man in diesem Kontext den Vergleich einer intuitiven Einschätzung auf Basis von „Thin slices“ mit einem Verhalten oder einem Ergebnis, das auf einer verwandten und dennoch anderen Skala gemessen wurde (Carney et al., 2007). In diesem Fall galt es, als Vergleichsmaßstab ein objektives Maß für Führungsstärke zu generieren. Führungsstärke als dispositionelles, relationales, aber vor allem soziales Merkmal manifestiert sich in einer konkreten Situation, in der jemand führt und andere folgen. Daher galt es eine Situation zu schaffen, in der tatsächliches Führungsverhalten beobachtbar und damit messbar wurde.

Zu diesem Zweck nahmen die Zielpersonen an einer Konstruktionsübung teil. Die Konstruktionsübung ist eine der Assessment Center-Praxis entnommene gruppendynamische „Übung, in der eine Gruppe von Teilnehmern gemeinsam eine praktische Aufgabe bearbeiten oder ein Produkt herstellen soll“ (Heilmann, 2002, S.105). Assessment-Center gehören zu den situativen beziehungsweise simulationsorientierten Verfahren der Personalauswahl und -entwicklung. Ihre Grundidee besteht im sich „Zeigen lassen“ und sie liefern somit die Möglichkeit einer Fremdbeobachtung von Verhalten und Kompetenzen. Die Teilnahme an einem Assessment Center kann die „Verhal-

tensklaviatur“ eines Teilnehmers aufzeigen. Es zeigt, ob ein Teilnehmer über die persönlichen Möglichkeiten verfügt, die ihn zur Führungsperson prädestinieren. Die Forschung zur prädiktiven Validität von Assessment Centern lieferte über die Jahre heterogene Ergebnisse. Aktuellere Übersichten beziffern sie derzeit mit etwa .37 (Höft & Funke, 2006). Obwohl sich die Assessment Center Praxis in den letzten Jahren stark verändert hat und weiterhin verändert, werden sie bis heute vornehmlich zur Rekrutierung von Führungskräften eingesetzt (Lievens & Thornton III, 2005).

Konstruktionsübungen eignen sich insbesondere als Beobachtungsgrundlage von Führungskompetenzen, sowie von sozialen und kommunikativen Fähigkeiten (Eilles-Matthiessen, al Hage, Janssen & Osterholz, 2002). Zu diesen Dimensionen gehören Eigenschaften wie Überzeugungskraft, Durchsetzungsvermögen, Kooperationsbereitschaft, usw. (Heilmann, 2002).

Als konkrete Aufgabenstellung wurde der Bau eines Turms gewählt. Da es sich bei dieser Übung um eine Gruppenübung handelt, die auch ohne konkreten berufs- oder firmenbezogenen Inhalt funktioniert, eignet sie sich sehr gut, um auch unter Studierenden Gruppen- und Führungsverhalten zu induzieren. Mit Hilfe verschiedener zur Verfügung stehender Werkzeuge, sollte innerhalb eines gegebenen Zeitrahmens von 20 Minuten ein Turm aus Pappe konstruiert werden, der bestimmten Kriterien genügt. Die Kriterien lauteten Höhe, Stabilität und Kreativität. Dieses Konstruktionsziel und der daran gekoppelte Wettbewerb mit den anderen teilnehmenden Gruppen, diente der Erhöhung der Seriosität der Aufgabe und der Motivation der Teilnehmer (Eilles-Matthiessen et al., 2002).

6.4.2. Methode

Zielpersonen

In einem ersten Schritt wurde das Stimulusmaterial für die Studie generiert. Für die Teilnahme an dieser Untersuchung konnten zu-

nächst in der Kölner Innenstadt 80 Personen gewonnen werden. Das Alter der Versuchspersonen variierte zwischen 18 und 37 Jahren, im Durchschnitt waren sie 25.05 Jahre alt ($SD = 3.62$). 41 Probanden waren männlichen (51.3%) und 39 weiblichen Geschlechts (48.8%). Im Versuchsraum stellten sich alle Teilnehmer einzeln etwa eine Minute in die laufende Videokamera vor. Der Bildausschnitt wurde dadurch standardisiert, dass alle Zielpersonen vor einem neutralen weißen Hintergrund an einem Tisch saßen und geradeaus in die Kamera blickten, so dass jeweils nur Gesicht und Oberkörper der Teilnehmer zu sehen waren.

Im Anschluss an die Einzelvideoaufnahmen nahmen die Zielpersonen an der Turmbauübung teil. Diese erfolgte jeweils in Gruppen von 4 Personen und wurde ebenfalls filmisch festgehalten. Aus Gründen der Vergleichbarkeit wurde die Geschlechterzusammensetzung der Gruppen konstant gehalten. Jede Gruppe bestand daher aus zwei weiblichen und zwei männlichen Teilnehmern. Ziel der Übung war es, innerhalb von 20 Minuten gemeinsam einen Turm aus Pappe zu bauen. Um die Motivation zu erhöhen und sicherzugehen, dass alle Teilnehmer die Aufgabe ernst nahmen, wurde ein Gewinnspiel ausgeschrieben, nach dem der beste Turm auf Basis der Kriterien Stabilität, Höhe und Originalität bewertet wurde (Heilmann, 2002). Der Preis für die Gewinnergruppe bestand in einer Flasche Wein und Kinogutscheinen für jede Person. Als Hilfsmittel standen Kleber, ein Lineal, Stifte und ein Konzeptpapier zur Verfügung.

Zum Abschluss der ersten Phase der Untersuchung wurden die Videosequenzen pro Zielperson jeweils auf 20 Sekunden gekürzt und zu einem tonlosen Gesamtfilm geschnitten. Einer Zielpersonen-Sequenz von 20 Sekunden folgte jeweils eine 9sekündige Pause, innerhalb derer der Bildschirm zunächst schwarz war und dann die Nummer des nächsten Filmausschnitts erschien. Die Gesamtdauer des Films mit allen 80 Zielpersonen betrug 40 Minuten.

Auf Basis der 20-sekündigen tonlosen Videosequenzen wurde die Attraktivität der Zielpersonen von 20 weiblichen und 20 männlichen Probanden bewertet. Diese waren im Durchschnitt 24.18 Jahre alt, wobei der jüngste 19 Jahre alt war und der älteste 32 ($SD = 2.90$). Ein Intraklassen-Koeffizient von .43 führte zu einem Cronbach's Alpha von .97, so dass die Attraktivitätsbewertungen zu einem Wert pro Zielperson aggregiert werden konnten. Die durchschnittliche Attraktivität der Zielpersonen lag bei $M = 3.34$ ($SD = 1.00$).

Die Verhaltensbeobachtung

Im Anschluss an die Generierung des Videomaterials begann die nächste Phase der Untersuchung, die der Erstellung des Referenzmaßes für Führungsstärke diente.

Entsprechend der Leitlinien der gängigen Assessment Center Praxis wurden zunächst führungsrelevante Verhaltensdimensionen definiert (Eilles-Matthiessen et al., 2002). Insgesamt wurden neun Kategorien aufgestellt, die jeweils mit konkreten Verhaltensankern versehen wurden. Somit erfolgte die Beobachtung des Führungsverhaltens anhand von klar definierten und kommunizierbaren Kriterien (Kleinmann, 2003), welche am aktuellen Forschungsstand zur charismatischen Führung (House & Howell, 1992; Weiss, 2003; Yukl, 2002) orientiert waren. Diese lauteten:

- Dimension: Aufgabeneinstieg

Beschreibung: Initiative, erste Dominanzhandlung in der neu gebildeten Gruppe.

Verhaltensanker: Person ergreift nach der Erklärung des Versuchsleiters als erste in der Gruppe das Wort.

- Dimension: aktive Beteiligung am Turmbau

Beschreibung: Person beteiligt sich tätlich am Turmbau.

Verhaltensanker: Person schneidet, klebt, dreht, hält fest, etc. Jede neue Verhaltensweise wird einfach dokumentiert.

- Dimension: Anzahl der eingebrachten Vorschläge und Lösungsansätze

Beschreibung: Person beteiligt sich verbal und visionär am Turmbau, macht generelle oder spezifische Vorschläge.

Verhaltensanker: „Ich finde, wir sollten...“, „Wir könnten...“, „Wenn wir... dann ginge auch...“

- Dimension: Lautstärke der Stimme

Beschreibung: Dominanz und Präsenz in der Gruppe. Wie laut spricht die Person im Durchschnitt im Vergleich zu den anderen Mitgliedern der Gruppe? Spricht sie leiser (-), gleich laut (0) oder lauter (+) als die anderen?

- Dimension: Anzahl der getroffenen Entscheidungen

Beschreibung: Wie dominant/durchsetzungsstark verhält sich die Person verbal und tätlich im Vergleich zu den anderen Gruppenmitgliedern?

Verhaltensanker: Person nimmt ohne vollständigen Konsens die Schere und schneidet einfach, klebt einfach, Person erstellt oder malt etwas (Fenster, Fahne, Blume, ...) ohne vorherige Absprache, „So, wir machen das jetzt einfach so...“, „Das geht nicht anders und deswegen wird das so gemacht.“

- Dimension: Verteilen von Aufgaben

Beschreibung: Person übernimmt eine Führungsrolle, delegiert Aufgaben.

Verhaltensanker: Person gibt einer anderen Person Schere, Kleber, Lineal, etc. verbunden mit einer konkreten Handlungsaufforderung

Als Moderationsfunktionen galten:

- Dimension: Berücksichtigung der verbleibenden Zeit

Beschreibung: Führungsfunktion durch Erinnerung der anderen Gruppenmitglieder an die Aufgabe und die verbleibende Zeit.

Verhaltensanker: Blick auf die Uhr, verbunden mit einer konkreten verbalen Äußerung über die Zeit: „Wir haben noch xx Minuten“, „Wir müssen uns etwas beeilen/schneller machen“

- Dimension: Ziel vor Augen

Beschreibung: Führungsfunktion durch Erinnerung der anderen Gruppenmitglieder an die verschiedenen Aspekte der Aufgabe

Verhaltensanker: Verbalisierung der verschiedenen Kriterien, Feststellung des Status Quo des geschaffenen Turms inklusive Diskrepanz zu bestimmten Kriterien, Beispiel: „Unser Turm ist jetzt sehr hoch, aber überhaupt nicht originell. Was können wir denn dafür noch machen?“

- Dimension: Motivation und Integration anderer

Beschreibung: Führungsfunktion durch Schaffen von Gruppenkohärenz und Zielorientierung

Verhaltensanker: aktives Fragen nach der Meinung der anderen, Aufforderung von Mitgliedern, die sich weniger beteiligen, Gruppenmitglieder loben

Die Auftretenshäufigkeit dieser als führungsstark definierten Verhaltensweisen wurde pro Zielperson gezählt und auf einem Beobachtungsbogen (Anhang D) dokumentiert. Vier geschulte Beobachter sahen in zwei Durchgängen alle 20-minütigen Videoaufnahmen aller 20 Versuchsgruppen. Bei jedem Durchgang konzentrierten sich jeweils zwei Beobachter auf das Verhalten einer Zielperson. Der Einsatz von zwei Beobachtern pro Zielperson wurde gewählt, um eventuelle Fehler und Verzerrungen gegenseitig auszugleichen.

Wie Tabelle 3 zu entnehmen ist, korrelierten die Beobachtungen der beiden Verhaltensauszähler auf allen neun Verhaltensdimensionen signifikant miteinander.

Da somit die Übereinstimmung der beiden Beurteiler bezogen auf die Ausführungshäufigkeit der einzelnen Handlungen als hoch anzusehen war, wurden alle Dimensionen zunächst beibehalten und es wurde für beide Beobachter der Mittelwert über alle Kategorien pro Zielperson berechnet.

Anschließend wurden die beiden durchschnittlichen Gesamtbeurteilungen zentriert. Es zeigte sich, dass auch die beiden Gesamturteile signifikant miteinander korrelierten ($r = .98$, $p < .001$). Es ist daher davon auszugehen, dass beide Beurteiler ihre Auszählung nach den gleichen Maßstäben durchführten.

Eine anschließende Reliabilitätsanalyse der Gesamtskala ergab ein Cronbach's Alpha von .75. Nach Ausschluss der Dimension Einstieg finden, die auf bivariater Ebene in keinem Zusammenhang mit der Gesamtskala stand ($r = .17$, n.s.), ergab sich ein zufrieden stellendes Cronbach's Alpha von .77, weshalb in einem finalen Schritt die Verhaltensbeobachtungsskala gebildet wurde, indem die beiden z-standardisierten Beobachtungen gemittelt wurden.

Im Folgenden wird daher der zentrierte Mittelwert aus beiden Verhaltensauszählungen als *Führungsverhalten* bezeichnet und dieser gilt als objektives Kriterium zur Feststellung der Führungsstärke der Zielpersonen.

Verhaltens- dimension	Interkorrela- tion der Verhaltens- auszähler	Mit- telwert Dimen- sion	Stan- dardab- weichung	Streuung	Korrela- tion mit Gesamt- skala
1. Aufgaben- einstieg	1.00**	.25	.44	0-1	.17
2. aktive Beteiligung	.93**	6.93	3.17	2-18	.60**
3. Vorschläge/ Lösungen	.95**	4.82	2.99	0-14.50	.82**
4. Lautstärke der Stimme	.98**	-.27	.59	-1-1	.65**
5. Anzahl Entscheidun- gen	.94**	.67	1.05	0-5.50	.50**
6. Verteilen von Aufgaben	.93**	.74	1.15	0-5	.76**
7. Berücksich- tigung der Zeit	.91**	.58	.85	0-3	.24*
8. Ziel vor Augen	.95**	1.80	2.31	0-15	.63**
9. Motivation/ Integration anderer	.91**	.90	1.52	0-10.50	.60**

Tabelle 2: Itemstatistik und Zusammenhänge innerhalb der Skala Führungsverhalten (*: $p < .05$; **: $p < .01$).

Beurteiler

64 Studenten verschiedener Fachrichtungen der Universität zu Köln beurteilten in Studie 2 die 80 Zielpersonen. Dabei wurde auf ein ausgewogenes Geschlechterverhältnis geachtet, so dass die Persönlichkeitsstärke von 34 weiblichen und 30 männlichen Probanden festgestellt wurde. Die Beurteiler befanden sich im Alter von 20 bis 29 Jahren, im Durchschnitt 23.27 Jahre ($SD = 2.37$).

Versuchsaufbau

In großen Teilen entsprach der Versuchsaufbau von Studie 2 dem von Studie 1. Alle intuitiven Beurteiler sahen im Labor den tonlosen Gesamtfilm mit allen 80 Zielpersonen. Der Versuchsleiter teilte den Probanden mit, dass es sich um eine Studie zum ersten Eindruck handele bei der es keine richtigen oder falschen Antworten gäbe, sondern lediglich ihr persönlicher Eindruck zähle. Daraufhin wurden die Fragebögen mit der Aufforderung ausgeteilt, die Instruktionen gut zu lesen, in denen die mündliche Instruktion wiederholt wurde. Nach jeder 20-sekündigen tonlosen Sequenz sollten die Probanden innerhalb von neun Sekunden ein Urteil über die Führungsstärke der soeben gesehenen Zielperson fällen. Diese wurde analog zu Studie 1 abgefragt (Anhang B). Nach der 40. Zielperson wurde eine Pause von zehn Minuten eingelegt.

Abschließend wurde den Teilnehmern gedankt und sie wurden über das Ziel der Studie aufgeklärt.

6.4.3. Ergebnisse

Zunächst wurden die Daten auf Beurteilerübereinstimmung hin überprüft. Hinsichtlich der Einschätzung von Persönlichkeitsstärke ergab sich ein Intraklassen-Koeffizient von .25, welcher zu einem Cronbach's Alpha von .96 führte. Die Interraterreliabilität war somit als ausreichend zu bewerten, so dass für die Folgeberechnungen die Einzelwerte zu einem Gesamtwert pro Zielperson aggregiert wurden.

Weiterhin wurden die Zielpersonen hinsichtlich ihrer Persönlichkeitsstärke systematisch unterschiedlich eingeschätzt. Die am wenigsten persönlichkeitsstark wahrgenommene Zielperson gelangte zu einem Durchschnitt von 2.33, die persönlichkeitsstärkste dagegen 5.40. Im Mittel erreichten die Zielpersonen eine zugeschriebene Persönlichkeitsstärke von 4.14 ($SD = .61$).

Im Folgenden standen die konkurrierenden Hypothesen zur Art und Weise des Einflusses der Zielpersonenmerkmale Geschlecht und Attraktivität im Zentrum der Analyse.

Beginnend mit der Attraktivität zeigte sich auf bivariater Ebene, dass die Attraktivität der Zielpersonen mit $r = .61$ signifikant mit der wahrgenommenen Persönlichkeitsstärke korrelierte ($p < .001$). Attraktiveren Zielpersonen wurde eine höhere Persönlichkeitsstärke zugeschrieben.

Der Zusammenhang zwischen Geschlecht und zugeschriebener Persönlichkeitsstärke wurde dagegen nicht signifikant ($r = -.08$, n.s.).

Daneben fand sich ebenfalls eine signifikante Korrelation zwischen Geschlecht und Attraktivität der Zielpersonen von $r = .26$ ($p < .05$). Da Frauen im Durchschnitt als attraktiver ($M = 3.60$, $SD = 1.04$) bewertet wurden als Männer ($M = 3.10$, $SD = .89$) und Attraktivität bei der Beurteilung der Persönlichkeitsstärke als Indikator verwendet wurde, führte dies zu ähnlich hohen Persönlichkeitsstärke-Mittelwerten auf bivariater Ebene.

Um die Einflüsse der beiden Prädiktorvariablen unabhängig von der Einwirkung der jeweils anderen zu untersuchen, wurde eine schrittweise Regression mit der intuitiv wahrgenommenen Persönlichkeitsstärke als abhängiger Variable und den beiden Faktoren Geschlecht und Attraktivität der Zielpersonen als unabhängigen Variablen im ersten Schritt durchgeführt. Bei der multivariaten Analyse verschwand der Suppressor-Effekt von Attraktivität und es zeigte sich, dass sowohl das Geschlecht der Zielpersonen ($\beta = .26$, $p < .05$) als auch die Attraktivität der Zielpersonen ($\beta = .67$, $p < .001$) von den

Intuitivbeurteilern als Indikatoren für Persönlichkeitsstärke genutzt wurden. Männlichen und attraktiveren Zielpersonen wurde jeweils eine höhere Persönlichkeitsstärke zugeschrieben. Durch Kenntnis dieser beiden Merkmale der Zielpersonen konnte etwa 43% der Varianz des Antwortverhaltens der Beurteiler erklärt werden.

Im zweiten Schritt der Regressionsanalyse wurde die Vorhersage des „Lack-of-fit“- Modells überprüft. Als dritte unabhängige Variable neben Geschlecht und Attraktivität wurde die Interaktion zwischen Geschlecht und Attraktivität aufgenommen. Zur Bildung des Interaktionsterms wurden zuvor beide Variablen zunächst z-standardisiert und anschließend miteinander multipliziert. Dieser dritte Einflussfaktor hatte allerdings keinerlei Auswirkung auf die Intuitivbeurteilung. Es resultierte keine Veränderung der erklärten Varianz des Gesamtmodells (Änderung in $R^2 = .00$, Änderung in $F = .00$, n.s.). Ein Zusammenwirken der beiden Merkmale Geschlecht und Attraktivität konnte somit nicht gefunden werden, weshalb die Vorhersage des „Lack-of-fit“- Modells auf Basis der vorliegenden Daten als nicht bestätigt anzusehen ist.

Im nächsten Schritt wurde ein mögliches Vorhandensein des „sexual attribution bias“ analysiert. In diesem Rahmen wurden zunächst Korrelationen zwischen der Attraktivität und der zugeschriebenen Persönlichkeitsstärke berechnet, jeweils getrennt sowohl nach dem Geschlecht der Zielperson als auch nach dem Geschlecht der Beurteiler. Weibliche Beurteiler nutzten bei ihrer Bewertung die Attraktivität als Indikator für Persönlichkeitsstärke. Dies galt sowohl bei der Einschätzung der weiblichen Zielpersonen ($r = .60$, $p < .001$) als auch bei männlichen Zielpersonen ($r = .65$, $p < .001$). Auch männliche Beurteiler zogen sowohl bei weiblichen ($r = .68$, $p < .001$) als auch bei männlichen Zielpersonen ($r = .61$, $p < .001$) deren Attraktivität als Bewertungsmaßstab heran. Ein Fisher Z-Test demonstrierte, dass sich die Stärke dieser Zusammenhänge nicht in Abhängigkeit des Geschlechts der Zielperson unterschied, weder bei männlichen Beurtei-

lern ($Z = -.55$ n.s.) noch bei weiblichen Beurteilern ($Z = .38$, n.s.). Es bleibt somit festzuhalten, dass sich männliche und weibliche Beurteiler sowohl bei der Beurteilung gleichgeschlechtlicher als auch bei andersgeschlechtlichen Zielpersonen in gleichem Maße an deren Attraktivität orientierten.

Als letzte Möglichkeit, ein Zuschreibungsmuster im Sinne eines „sexual attribution bias“ zu finden, wurde das Vorliegen von kurvilinearen Zusammenhängen geprüft. Es wurden vier Kurvenanpassungen berechnet, die Attraktivität der Zielpersonen als unabhängige Variable und die Urteile über deren Persönlichkeitsstärke als abhängige Variable. So resultierten die vier Kombinationen (1) Einschätzung der weiblichen Zielpersonen durch weibliche Beurteiler, (2) durch männliche Beurteiler, (3) Einschätzung der männlichen Zielpersonen durch weibliche Beurteiler und (4) durch männliche Beurteiler.

Tabelle 3 gibt die aufgeklärte Varianz der vier Modelle, jeweils unter der Annahme eines linearen beziehungsweise quadratischen Zusammenhangs, an. Ein Steigers Z-Test verdeutlichte, dass in keiner der vier Kombinationen die Modellierung eines quadratischen Zusammenhangs zu einem Anstieg der erklärten Varianz führte.

Zusammenhang	<i>weibliche Beurteiler, weibliche Zielpersonen</i>	<i>Weibliche Beurteiler, männliche Zielpersonen</i>	<i>Männliche Beurteiler, weibliche Zielpersonen</i>	<i>Männliche Beurteiler, männliche Zielpersonen</i>
Linear	36%	43%	46%	37%
Quadratisch	39%	45%	49%	37%
Steigers Z	-1.16, n.s.	-1.07, n.s.	-1.18, n.s.	.00, n.s.

Tabelle 3. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielperson (*: $p < .05$; **: $p < .01$).

Als Zwischenfazit ist festzuhalten, dass von den drei konkurrierenden Hypothesen die Nutzung des Geschlechts- sowie des Attraktivitätsstereotyps bestätigt wurde. Die Vorhersagen des „Lack-of-fit“- Modells sowie des „sexual attribution bias“ konnten nicht belegt werden.

Das Vorliegen von Genauigkeit bei der Wahrnehmung von Führungsstärke wurde vorab als Korrespondenz zwischen intuitiv zugeschriebener Führungsstärke und der Skala Führungsverhalten definiert. Die Untersuchung dieser Frage ergab eine signifikante Korrelation zwischen der intuitiv wahrgenommenen Persönlichkeitsstärke und der Skala Führungsverhalten ($r = .23$, $p < .05$). Die Beurteiler konnten die Führungsstärke der Zielpersonen besser als zufällig erkennen.

Auf multivariater Ebene konnte dieses erste Resultat jedoch nicht repliziert werden. In einer schrittweisen Regression wurden im ersten Schritt die Einflüsse der unabhängigen Variablen Geschlecht, Attraktivität und der Skala Führungsverhalten auf die abhängige Variable zugeschriebene Persönlichkeitsstärke untersucht. Tabelle 4 zeigt, dass über die beiden signifikanten Kriterien Geschlecht ($\beta = -.23$, $p < .05$) und Attraktivität ($\beta = .65$, $p < .001$) hinaus die Skala Führungsverhalten keinen Einfluss auf die zugeschriebene Persönlichkeitsstärke hatte.

Variable	Modell 1	Modell 2
Führungsverhalten	.13	.13
Geschlecht	-.23*	-.23*
Attraktivität	.65**	.65**
Interaktionsterm		.01
R^2	.45**	.45**

Tabelle 4. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Führungsverhalten, Geschlecht und Attraktivität sowie des Interaktionsterms auf die zugeschriebene Persönlichkeitsstärke (*: $p < .05$; **: $p < .01$).

Der Vollständigkeit halber wurde im zweiten Schritt die Interaktion zwischen Geschlecht und Attraktivität auf ihren Aufklärungswert hin überprüft und als vierte unabhängige Variable in die Gleichung aufgenommen. Auch hier änderte sich die erklärte Varianz durch die Aufnahme des Interaktionsterms nicht (Änderung in $F = .00$, Änderung in $R^2 = .00$, n.s.). Die Beurteiler orientierten sich bei der Einschätzung der Persönlichkeitsstärke somit an dem Geschlecht und der Attraktivität der Zielpersonen und erkannten darüber hinaus nicht die wahre Ausprägung des Merkmals.

Um abschließend die „Körnchen Wahrheit“- Hypothese zu überprüfen, wurde eine schrittweise Regression mit den unabhängigen Variablen Geschlecht und Attraktivität der Zielpersonen und als abhängiger Variable die Skala Führungsverhalten durchgeführt. So konnte überprüft werden, ob Geschlecht und Attraktivität valide Hinweise auf tatsächlich gezeigtes Führungsverhalten darstellen.

Tabelle 5 zeigt, dass ein Einfluss der genannten Art nicht vorlag. Weder das Geschlecht ($\beta = -.16$, n.s.), noch die Attraktivität ($\beta = .15$, n.s.) der Zielpersonen, noch der im zweiten Schritt als weitere unabhängige Variable hinzugenommene Interaktionsterm (Änderung in $R^2 = .01$, Änderung in $F = .77$, n.s.) hatte eine ursächliche Wirkung auf das tatsächlich gezeigte Führungsverhalten.

Variable	Modell 1	Modell 2
Geschlecht	-.16	-.16
Attraktivität	.15	.16
Interaktionsterm		-.10
R^2	.04**	.05**

Tabelle 5. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie deren Interaktion auf die Skala Führungsverhalten (*: $p < .05$; **: $p < .01$).

Somit lässt sich festhalten, dass auf Basis der vorliegenden Daten kein „Körnchen Wahrheit“ im Geschlechts- und Attraktivitätsstereotyp gefunden werden konnte.

6.4.4. Diskussion

Die Durchführung von Studie 2 geschah mit zwei zentralen Zielen. Nachdem Studie 1 bereits gezeigt hatte, dass Intuitivbeurteiler für ihr Urteil sowohl das Attraktivitäts- als auch das Geschlechtsstereotyp auf direktem Wege heranziehen, sollte dieser Effekt in Studie 2 repliziert werden.

Es zeigte sich auch in dieser Studie, dass der Prozess der Zuschreibung von Führungsstärke direkt jeweils vom „what is beautiful is good“- (Dion et al., 1972; Langlois et al., 2000) und vom „think manager - think male“- Stereotyp (Schein, 1973, 1975; Sczesny, 2003b; Sczesny et al., 2006) beeinflusst wurde. Unter den Zielpersonen wurde jeweils den attraktiveren sowie insgesamt den Männern eine höhere Führungsstärke attestiert. Da die weiblichen Zielpersonen sich im Durchschnitt durch eine höhere Attraktivität auszeichneten, blieb der Einfluss des Geschlechterstereotyps auf bivariater Ebene zunächst verborgen. Erst auf multivariater Ebene verschwand der Suppressor-Effekt.

Wie bereits in Studie 1 gesehen, konnten dagegen die Hypothesen des „Lack-of-fit“- Modells (Heilman, 1983) sowie des „sexual attribution bias“ (Försterling et al., 2007) nicht bestätigt werden. Bezogen auf das „Lack-of-fit“- Modell hatte die Interaktion zwischen Geschlecht und Attraktivität keinen Einfluss auf das Antwortverhalten der Intuitivbeurteiler. Bezogen auf die Annahmen des „sexual attribution bias“ zeigte sich bei der gleichgeschlechtlichen Beurteilung kein kurvilinearere Zusammenhang zwischen der Attraktivität der Zielpersonen und der ihnen zugeschriebenen Führungsstärke. Unabhängig von dem Geschlecht des Beurteilers und dem Geschlecht der Zielperson erwies sich jeweils der lineare Zusammenhang als angemessen.

Neben der Analyse dieser Zuschreibungsprozesse sollte Studie 2 jedoch vorrangig Aufschluss über die Frage geben, wie gut Führungsstärke intuitiv erkennbar ist.

Zu diesem Zweck wurde zunächst ein geeignetes Maß zur Feststellung von Genauigkeit definiert. Als Indikator für Genauigkeit wurde eine substanzielle Beziehung zwischen den intuitiven Einschätzungen von Führungsstärke und der tatsächlichen Führungsstärke der Zielpersonen festgelegt (Funder, 1999). Die tatsächliche Führungsstärke wurde anhand einer Auszählung führungsrelevanter Verhaltensweisen während einer Assessment Center Übung objektiviert. Die Festlegung der führungsstarken Verhaltensweisen erfolgte auf der Basis der aktuellen Forschung über charismatische Führung (House & Howell, 1992; Weiss, 2003; Yukl, 2002).

Auch wenn die soziale Kognitionsforschung kaum Aussagen über die Genauigkeit von Personenwahrnehmung trifft, war aus der Hypothese des Persönlichkeitseinschätzungs-Instinkts (Haselton & Funder, 2006) heraus vermutet worden, dass die Führungsstärke fremder Personen rasch und besser als zufällig eingeschätzt werden kann. Führungsstärke ist eng mit dem sozialen Status einer Person innerhalb einer Gruppe verknüpft und da Gruppenzugehörigkeit und der relative soziale Status unmittelbare Auswirkungen auf das Überleben und die Reproduktion haben, sollten Menschen als Grundlage für angemessenes Folgeverhalten die Fähigkeit besitzen, andere Menschen in dieser Hinsicht schnell und zuverlässig einzuschätzen.

Mit Hilfe eines „Thin slices“- Versuchsaufbaus konnte diese Hypothese zunächst bestätigt werden. Auf bivariater Ebene zeigte sich ein signifikanter Zusammenhang zwischen tatsächlicher und intuitiv vermuteter Führungsstärke. Würde man auf dieser Analyseebene verbleiben, so könnte dieses Ergebnis als Beleg für die Hypothese gedeutet werden, dass Menschen von Natur aus generell valide Urteile über andere Menschen fällen können.

Allerdings zeigten weitere Analysen, dass der gefundene Effekt von .23 auf bivariater Ebene auf .13 auf multivariater Ebene schrumpfte und hier lediglich 2% des Antwortverhaltens der Beurteiler erklären konnte. Ein Grund für die geringe Effektstärke ist, dass sich die Beurteiler auf Stereotype verließen. So konnte die Attraktivität der Zielpersonen insgesamt etwa 45% des Antwortverhaltens der Beurteiler erklären und das Geschlecht der Zielpersonen etwa 7%. Die wahre Führungsstärke der Zielpersonen konnte von den Beurteilern über die Stereotype hinaus nicht akkurat erkannt werden.

Nun kann die Orientierung an wenigen Merkmalen grundsätzlich dann eine sinnvolle Herangehensweise darstellen, wenn als Bewertungsgrundlage valide Hinweise auf das in Frage stehende Merkmal genutzt werden. Diese Herangehensweise entspräche einer „Take-the-best“-Heuristik im Sinne einer simplen und sparsamen Heuristik (Gigerenzer & Goldstein, 2001). Um dies zu überprüfen, wurden die beiden Stereotype hinsichtlich ihres „Körnchen Wahrheits“ untersucht. Allerdings zeigte sich keine Entsprechung von Geschlecht und Attraktivität in realem Führungsverhalten. Weder zeigten Männer mehr Führungsverhalten noch allgemein attraktivere Teilnehmer.

Ein möglicher Grund, warum die wahre Führungsstärke nicht zutreffend erkannt werden konnte, könnte in dem Konstrukt Führungsstärke selber liegen. So ist es denkbar, dass Studenten, die noch wenig Erfahrung mit Führung und Arbeit in Gruppen haben, noch keine implizite Definition von Führungsstärke ausgebildet haben. Dementsprechend wäre es möglich, dass sie aus diesem Grund nicht wissen wie und woran Führungsstärke erkennbar ist. Aus diesem Mangel an validen Hinweisen könnte die Orientierung an salienten Stereotypen resultieren. Möglich wäre es daher, dass die Betonung des Geschlechts und der Attraktivität der zu beurteilenden Personen im Laufe des Lebens und mit steigender Erfahrung abnimmt. Der Überprüfung dieser Möglichkeit gilt Studie 3.

6.5. Studie 3

Studie 3 wurde mit dem vorrangigen Ziel durchgeführt, die Rolle von Expertise und Erfahrung im intuitiven Beurteilungsprozess zu untersuchen. Der Unterschied dieser im Vergleich zu den beiden vorherigen Studien lag in der Zusammensetzung der Beurteilerstichprobe. Die gleichen Zielpersonen von Studie 2 wurden in Studie 3 einerseits von einer studentischen Stichprobe eingeschätzt und andererseits von einer Gruppe von Beurteilern, deren Beruf regelmäßig die Auswahl von Personal beinhaltet. Aufgrund der unterschiedlichen Erfahrungen mit Beurteilungssituationen werden diese beiden Gruppen im Folgenden als *Laien* beziehungsweise *Experten* bezeichnet.

Zum einen sollte mit Hilfe dieses Versuchsaufbaus der Frage nachgegangen werden, ob Menschen die Führungsstärke anderer zutreffender einschätzen können, wenn sie mehr berufsbedingte Erfahrung sowohl mit dem Merkmal Führungsstärke als auch mit dem Treffen von Auswahlentscheidungen haben. Mit anderen Worten wurde untersucht, ob sich Menschenkenntnis im Laufe des Lebens verbessert.

Zum anderen sollte erforscht werden, ob die Merkmale Attraktivität und Geschlecht mit steigender Expertise in geringerem Umfang genutzt werden. Möglicherweise werden im Laufe des Lebens andere - bessere? - Hinweise auf Führungsstärke gelernt und angewendet. So wäre es denkbar, dass Menschen sich vor allem aus Mangel an besseren Bewertungsmaßstäben auf Stereotype als Heuristiken verlassen.

Der Vergleich der Urteilsgenauigkeit von Laien mit Experten dient weiterhin der Beleuchtung eines übergeordneten Aspekts der Fragestellung. Folgt man der Annahme eines Persönlichkeitseinschätzungs-Instinktes (Haselton & Funder, 2006), so stellt die soziale Wahrnehmungsfähigkeit des Menschen eine so grundlegende Fähigkeit dar, dass sie auch ohne explizites Training auftreten sollte. In

diesem Falle sollte mit steigender Expertise keine wesentliche Verbesserung der Einschätzungsgüte zu verzeichnen sein.

Auf der anderen Seite werden Heuristiken zur Komplexitätsreduktion in Situationen unter Ungewissheit angewendet (Gilovich et al., 2002). Je mehr Wissen eine Person auf einem speziellen Gebiet hat, desto weniger anfällig sollte sie für verzerrende Einflüsse sein. Insofern scheint es plausibel, dass Laien, die noch kein Wissen über adäquates Führungsverhalten erworben haben, sich eher auf Heuristiken verlassen als Experten. Experten dagegen verfügen berufsbedingt über mehr Erfahrung mit dem Erkennen von Führungsstärke und hatten daher die Möglichkeit, sich auf diesem Gebiet eine gewisse Expertise anzueignen.

Diese Argumentation entspricht der Lernperspektive der Intuition. Nach Plessner und Czenna (2007) arbeitet das intuitive oder spontane Urteilssystem vor allem auf Basis von implizitem Wissen und Einstellungen, die automatisch aktiviert werden, das Verhalten steuern und durch assoziatives Lernen ausgebildet werden. Implizite Einstellungen sind damit konträr zu expliziten Einstellungen, die auf Basis der in einem Moment zur Verfügung stehenden Informationen entstehen.

Aus der Lernperspektive heraus ist Intuition als eine Form von Expertise anzusehen (Hogarth, 2001), als ein Prozess, der auf Basis von im Langzeitgedächtnis gespeichertem Wissen operiert (Plessner, Betsch & Betsch, 2007). Auch Klein (2004) definiert Intuition als „the way we translate our experiences into action“ (S. 4). Nach seiner Auffassung führt wiederholtes Erleben bestimmter Situationen zu einem insgesamt größeren Repertoire an Verhaltensmustern, Mechanismen und Wissensstrukturen. Experten erleben neue Situationen somit nicht mehr als neuartig, sondern sie erkennen Parallelen zu vorherigen Situationen und wissen daher automatisch und intuitiv, wie sie sich zu verhalten haben.

So könnten Experten bei Bewerbern beispielsweise intuitiv bestimmte Fähigkeitskomponenten oder Kombinationen bestimmter Fähigkeiten wahrnehmen, die erfahrungsgemäß mit hohen Führungsqualitäten einhergehen. Folgt man somit der Lernperspektive der Intuition (Klein, 1998, 2004; Plessner et al., 2007), sollten erfahrene Personenbeurteiler im Vergleich zu Laien die besseren Einschätzungen von Führungskraft abgeben.

Zur Ausbildung von Expertise sind zwei Voraussetzungen notwendig, Wiederholung und Feedback (Klein, 2004). Experten benötigen viel Übung, um mit den jeweiligen Situationen vertraut zu werden und ihre eigenen Verhaltensroutinen zu entwickeln, das heißt ihre optimale Strategie zu finden, um gute Entscheidungen zu treffen (Betsch, Brinkmann, Fiedler & Breining, 1999). Aus diesem Grund wurde bei der Rekrutierung der Expertenstichprobe für Studie 3 darauf geachtet, dass die Teilnehmer seit mindestens einem Jahr im Unternehmen an der Personalbeschaffung beteiligt sind.

Die zweite wichtige Voraussetzung zur Entstehung von Expertise ist Feedback. Experten sollten regelmäßige Rückmeldungen über die Güte ihrer Entscheidung erhalten, indem sie über die Leistung der von ihnen ausgewählten Mitarbeiter informiert werden. Nur durch einen qualitativ hochwertigen wiederkehrenden Vergleich ihrer Entscheidungen mit der Realität können Experten im Laufe der Zeit kumulativ an der Realität getestete Beurteilungsstrategien entwickeln, die sie zu intuitiv besseren Urteilen im Vergleich zu Laien befähigen (Hogarth, 2001, 2005; Klein, 2004). Aus diesem Grund wurde bei den Experten zusätzlich erhoben, ob diese Feedback über die Leistung der von ihnen ausgewählten Mitarbeiter erhalten.

6.5.1. Methode

Zielpersonen und Verhaltensbeobachtung

Als Stimulusmaterial fungierten in dieser Studie die gleichen, bereits in Studie 2 verwendeten Videosequenzen. Als Referenzmaß zur Fest-

stellung der Genauigkeit der Einschätzungen wurde die Skala Führungsverhalten herangezogen, welche ebenfalls im Rahmen von Studie 2 generiert wurde. Diese basierte auf einer kriterienorientierten Auszählung des von den Zielpersonen in einer Assessment Center Übung gezeigten Führungsverhaltens.

Da die professionellen Beurteiler über ein sehr eingeschränktes Zeitbudget verfügten, wurde aus den 80 in Studie 2 verwendeten Zielpersonen eine zufällige Stichprobe von 40 Zielpersonen gezogen. 20 der zufällig ausgewählten Zielpersonen waren weiblichen Geschlechts, die anderen 20 waren männlich. Ihr Alter variierte zwischen 19 und 37 Jahren, im Durchschnitt lag ihr Alter bei 24.58 Jahren ($SD = 3.49$).

Hinsichtlich ihrer Attraktivität rangierten die 40 Zielpersonen in der Wahrnehmung der Beurteiler zwischen 1.50 und 5.65, im Durchschnitt 3.28 ($SD = 1.03$). Diese war bereits in Studie 2 von insgesamt 20 männlichen und 20 weiblichen Studenten der Universität zu Köln im Alter von 19 bis 32 Jahren, im Durchschnitt 24.18 Jahre ($SD = 2.90$), festgestellt worden ($ICC = .43$, Cronbach's $\alpha = .97$).

Beurteiler

Die Persönlichkeitsstärke der Zielpersonen wurde durch zwei Gruppen von Beurteilern eingeschätzt. Die Gruppe der *Laien* bestand aus 42 studentischen Beurteilern mit einem Durchschnittsalter von 24.26, wobei die jüngste Versuchsperson 19 Jahre und die älteste 28 Jahre alt war ($SD = 2.42$). 21 von ihnen waren weiblich und weitere 21 waren männlich. Die Gruppe der *Experten* bestand aus 42 professionellen Beurteilern, davon 22 Frauen und 20 Männer, im Alter von 27 bis 63 Jahren, im Durchschnitt 40.62 Jahre ($SD = 9.65$). Hinsichtlich ihrer Qualifikation gaben 22 der 42 Experten an, ein Studium beendet zu haben, 20 hatten eine Ausbildung absolviert.

Eine Bedingung der Experten für die Teilnahme an dieser Untersuchung war die Erfahrung mit Personalauswahlentscheidungen. Als Kriterium wurde zuvor festgelegt, dass alle Experten seit mindestens

einem Jahr am Prozess der Personalauswahl im Unternehmen beteiligt sein müssen. Aufgrund dieses Kriteriums trafen die Personen somit seit mindestens einem Jahr Auswahlentscheidungen, wobei die erfahrenste Person bereits eine 40jährige Erfahrung in der Personalauswahl aufwies. Die Standardabweichung betrug 8.99 Jahre und im Durchschnitt trafen die Experten seit 10.02 Jahren regelmäßig Personalauswahlentscheidungen. 25 von ihnen waren dabei direkt in der Personalabteilung ihres Unternehmens beschäftigt, 9 waren als Abteilungsleiter an der Personalbeschaffung für die ihnen unterstellten Teams beteiligt und die 8 restlichen Experten entstammten verschiedenen Geschäftsbereichen.

Darüber hinaus zeigte sich, dass die rekrutierten Experten sehr rege an der Personalauswahl beteiligt waren. In Bezug auf die Regelmäßigkeit dieser Aufgabe machten 7 Experten keinerlei Angaben, während die anderen pro Monat zwischen 1 und 120 Entscheidungen angaben, im Durchschnitt 32.51 ($SD = 43.21$).

Bezogen auf die Frage nach der Feedbackpraxis im Unternehmen gaben insgesamt 22 Personen an, regelmäßige Rückmeldungen über die Leistung der von ihnen ausgewählten Kandidaten zu erhalten, während 15 kein Feedback bekamen.

Versuchsaufbau

Bezogen auf die studentischen Beurteiler entsprach der Versuchsaufbau dieser Studie denen der ersten beiden Studien. In Räumen der Universität zu Köln fanden mehrere Versuchstermine statt, zu welchen die Versuchspersonen als Beurteiler rekrutiert wurden. Dort sahen sie die 40 tonlosen Videosequenzen der Zielpersonen als Gesamtfilm. Die Beurteiler erhielten sowohl mündlich als auch schriftlich die Information, dass es sich um eine Untersuchung zum Thema „erster Eindruck“ handele, im Zuge derer es keine richtigen oder falschen Antworten gäbe. Die Beurteiler wurden gebeten, nach jeder 20-sekündigen tonlosen Sequenz innerhalb der Zeit zwischen den einzelnen Szenen (neun Sekunden) ein Urteil über die Zielperso-

nen anhand der drei Fragen der Skala Persönlichkeitsstärke nach Noelle-Neumann abzugeben. Die intuitiv zugeschriebene Führungsstärke wurde somit analog zu den Studien 1 und 2 erhoben (Anhang B). Nach Ende des Films bedankte sich der Versuchsleiter bei den Probanden und klärte sie über das Ziel der Studie auf.

Hinsichtlich der Instruktionen und des allgemeinen Ablaufs unterschied sich der Versuchsaufbau bei den professionellen Beurteilern nicht von der Befragung der Laien. Anders als bei den studentischen Beurteilern, fand der Versuch bei den Experten in zu diesem Zweck zur Verfügung gestellten Räumen der jeweiligen Unternehmen statt.

6.5.2. Ergebnisse

Der erste Analyseschritt galt der Überprüfung der Interraterreliabilität beider Beurteilergruppen. Sowohl bei den Laien als auch bei den Experten konnten die Persönlichkeitsstärkeeinschätzungen zu einem Gesamtmittelwert pro Zielperson aggregiert werden. Bei den Laien führte ein Intraklassen-Koeffizient von .25 zu einem Cronbach's Alpha von .93, während bei den Experten das Cronbach's Alpha von .94 auf einem Intraklassen-Koeffizient von .26 beruhte.

Laien und Experten nahmen die Persönlichkeitsstärke der 40 Zielpersonen systematisch unterschiedlich wahr. Die Laien schrieben den Zielpersonen im Durchschnitt eine Persönlichkeitsstärke von 4.26 ($SD = .63$) zu und die Experten durchschnittlich 4.08 ($SD = .62$). Die am wenigsten persönlichkeitsstark wahrgenommene Person war bei beiden Beurteilergruppen die gleiche. Sie erhielt von den Laien eine durchschnittliche Bewertung von 2.67 und von den Experten 2.25. Auch die Zielperson, der die höchste Persönlichkeitsstärke zugeschrieben wurde, war bei beiden Gruppen die gleiche. Sie wurde von den Laien im Mittel mit 5.59 und von den Experten mit 5.39 bewertet.

Nachdem festgestellt worden war, dass beide Gruppen von Beurteilern erstens die gleiche Dimension beurteilten und zweitens die Beur-

teilungen systematisch variierten, wurde sich zunächst der Frage nach dem intuitiven Beurteilungsprozess zugewandt.

Bezogen auf die Nutzung des Attraktivitätsstereotyps zeigte sich, dass beide Gruppen die Attraktivität der Zielpersonen als Hinweis auf deren Führungsstärke nutzten. Die Korrelation zwischen Attraktivität der Zielpersonen und Intuitivurteil der Laien lag bei $r = .63$ ($p < .001$) und unterschied sich damit nicht ($Z = .23$, n.s.) von dem Zusammenhang bei den Experten mit $r = .66$ ($p < .001$).

Die Korrelation zwischen Geschlecht und den Intuitivurteilen wurde dagegen weder bei den Experten ($r = -.10$, n.s.) noch bei den Laien ($r = -.08$, n.s.) signifikant. Entgegen der in den Vorgängerstudien gefundenen Korrelation zwischen Attraktivität und Geschlecht der Zielpersonen, wurde diese hier nicht signifikant ($r = .18$, n.s.).

Anschließend wurde für beide Beurteilergruppen jeweils getrennt eine simultane Regression mit der abhängigen Variablen zugeschriebene Persönlichkeitsstärke und den beiden unabhängigen Variablen Geschlecht und Attraktivität im ersten Schritt berechnet.

Beginnend mit dem Intuitivurteil der Laien zeigte sich, dass Attraktivität mit einem β von $.67$ ($p < .001$) einen signifikanten Einfluss auf die intuitive Bewertung besaß, das Geschlecht der Zielpersonen jedoch nicht ($\beta = -.20$, n.s.). Bezogen auf das Intuitivurteil der Experten zeigte sich ein signifikanter Einfluss der Attraktivität der Zielpersonen ($\beta = .70$, $p < .001$), sowie ein marginal signifikanter Einfluss des Geschlechts der Zielpersonen ($\beta = -.23$, $p < .07$).

Es lässt sich somit festhalten, dass sich ein direkter Einfluss des Attraktivitätsstereotyps in beiden Gruppen von Beurteilern finden ließ, während das Geschlecht der Zielpersonen einen lediglich marginal signifikanten Einfluss auf das Urteil der Experten besaß.

Zur Überprüfung der Annahme des „Lack-of-fit“- Modells wurde im zweiten Schritt der Regression bei beiden Beurteilergruppen der Interaktionsterm als dritte unabhängige Variable aufgenommen.

Dieser war zuvor aus der Multiplikation der zentrierten Attraktivität und dem Geschlecht der Zielpersonen gebildet worden. Weder bei den Laien (Änderung in $F = .35$, Änderung in $R^2 = .01$, n.s.) noch bei den Experten (Änderung in $F = .32$, Änderung in $R^2 = .01$, n.s.) führt die Aufnahme des Interaktionsterms jedoch zu einer signifikanten Verbesserung der Modellpassung. Ein Zusammenwirken der beiden Merkmale Geschlecht und Attraktivität konnte somit in der vorliegenden Studie nicht bestätigt werden.

Bezogen auf den „sexual attribution bias“ wurde in den beiden vorhergehenden Studien gezeigt, dass das postulierte Attributionsmuster zumindest bei studentischen Beurteilern nicht vorlag. Auch in dieser Studie wurde diese Frage überprüft. Zunächst wurden in beiden Beurteilungsgruppen für die weiblichen und die männlichen Beurteiler getrennt nach dem Geschlecht der Zielperson je eine Korrelation zwischen der Attraktivität der Zielpersonen und der ihnen zugeschriebenen Persönlichkeitsstärke berechnet.

Bei den Laien zeigte sich bei der Beurteilung des anderen Geschlechts jeweils ein positiver Zusammenhang, der sich bei der Bewertung von männlichen Zielpersonen durch weibliche Beurteiler ($r = .70$, $p < .001$) nicht signifikant ($Z = 1.10$, n.s.) von der Beurteilung von weiblichen Zielpersonen durch männliche Beurteiler ($r = .55$, $p < .05$) unterschied. Beide Geschlechter sprachen den attraktiveren Vertretern des anderen Geschlechts eine höhere Persönlichkeitsstärke zu. Die Korrelation zwischen der Attraktivität der Zielpersonen und der wahrgenommenen Persönlichkeitsstärke beim jeweils gleichen Geschlecht war ebenfalls in beiden Fällen signifikant positiv. Sie lag bei männlichen Beurteilern und männlichen Zielpersonen bei $r = .75$ ($p < .001$) und war damit nicht signifikant stärker ausgeprägt ($Z = 1.44$, n.s.) als zwischen weiblichen Beurteilern und weiblichen Zielpersonen ($r = .57$, $p < .01$).

Auch bei der Beurteilung durch die Experten waren alle Zusammenhänge positiv und signifikant, sowohl zwischen der Attraktivität der

gegengeschlechtlichen Zielpersonen und der zugeschriebenen Persönlichkeitsstärke durch weibliche Beurteiler ($r = .79$, $p < .001$) als auch durch männliche Beurteiler ($r = .57$, $p < .01$). Der Zusammenhang war bei weiblichen Zielpersonen marginal signifikant stärker ($Z = 1.87$, $p < .07$) ausgeprägt als bei männlichen Experten und den männlichen Zielpersonen. Bei der gleichgeschlechtlichen Bewertung fand sich kein signifikant unterschiedlich ausgeprägter Zusammenhang ($Z = .58$, n.s.). Die Korrelation lag bei weiblichen Beurteilern und weiblichen Zielpersonen bei $r = .66$ ($p < .01$) und zwischen männlichen Beurteilern und männlichen Zielpersonen bei $r = .58$ ($p < .01$).

Auch hier wurde in einem abschließenden Schritt die Adäquatheit der Annahme eines linearen Zusammenhanges im Vergleich zu einem kurvilinearen Zusammenhang untersucht. Tabellen 6 und 7 zeigen die erklärte Varianz der jeweils vier berechneten Kurvenanpassungen pro Gruppe.

Zusammenhang	<i>weibliche Beurteiler, weibliche Zielpersonen</i>	<i>Weibliche Beurteiler, männliche Zielpersonen</i>	<i>Männliche Beurteiler, weibliche Zielpersonen</i>	<i>Männliche Beurteiler, männliche Zielpersonen</i>
Linear	32%	48%	30%	56%
Quadratisch	42%	53%	41%	57%
Steigers Z	-2.87**	-1.46	-3.27**	-.40

Tabelle 6. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhanges zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke bei Laien getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielpersonen (*: $p < .05$; **: $p < .01$).

Bei den Laien zeigt sich, dass der Zusammenhang zwischen der Attraktivität der Zielpersonen und der zugeschriebenen Persönlichkeitsstärke in den Beurteilungskombinationen (1) Einschätzung der

weiblichen Zielpersonen durch weibliche Beurteiler und (3) Einschätzung der männlichen Zielpersonen durch weibliche Beurteiler besser durch einen quadratischen als durch einen linearen Zusammenhang modelliert werden konnte. In den beiden anderen Kombinationen erwies sich der lineare Zusammenhang als ebenso angemessen.

Auch bei den Experten konnte die Kombination (3) Einschätzung der männlichen Zielpersonen durch weibliche Beurteiler adäquater durch die Annahme eines quadratischen Zusammenhangs abgebildet werden. In allen anderen Fällen führte die Annahme eines kurvilinearen gegenüber eines quadratischen Zusammenhangs nicht zu einem signifikanten Anstieg der erklärten Varianz.

Zusammenhang	<i>weibliche Beurteiler, weibliche Zielpersonen</i>	<i>Weibliche Beurteiler, männliche Zielpersonen</i>	<i>Männliche Beurteiler, weibliche Zielpersonen</i>	<i>Männliche Beurteiler, männliche Zielpersonen</i>
Linear	43%	62%	33%	34%
Quadratisch	48%	68%	40%	40%
Steigers Z	-1.31	1.58	-1.97*	-1.95

Tabelle 7. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke bei Experten getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielpersonen (*: $p < .05$; **: $p < .01$).

Somit bleibt festzuhalten, dass ein „sexual attribution bias“ weder bei den Laien noch bei den Experten vorlag. Ein kurvilinear-er gleichgeschlechtlicher Zusammenhang war bei Laien lediglich in einem Fall zu finden, bei den Experten fand sich keiner. Die beiden anderen gefundenen kurvilinearen Zusammenhänge bestanden bei der gegengeschlechtlichen Beurteilung von weiblichen Zielpersonen durch männliche studentische sowie durch professionelle Beurteiler. In

beiden Fällen hätte der „sexual attribution bias“ jedoch einen linearen Zusammenhang vorhergesagt.

Im Folgenden wurde der Frage nachgegangen, ob die Laien beziehungsweise die Experten die Führungsstärke der Zielpersonen erkennen konnten. Zunächst wurden jeweils Korrelationen zwischen der aggregierten Intuitivbewertung und der Skala Führungsverhalten berechnet. Es zeigte sich, dass beide Korrelationen signifikant waren, bei den Laien mit einem Korrelationskoeffizient von $r = .28$ ($p < .05$) und bei den Experten mit $r = .29$ ($p < .05$). Die Experten schätzten die Zielpersonen somit nicht besser ein als die Laien ($Z = .05$, n.s.). Im Gegenteil, das Urteil der Experten korrespondierte in hohem Maße mit der Einschätzung der Laien ($r = .91$, $p < .001$).

Zur weiteren Analyse dieses Zusammenhangs wurden getrennt für die beiden Beurteilergruppen jeweils schrittweise Regressionen mit der abhängigen Variablen Intuitivurteil und den unabhängigen Variablen Attraktivität und Geschlecht der Zielpersonen im ersten Schritt und der Hinzunahme des Interaktionsterms aus Attraktivität und Geschlecht im zweiten Schritt gerechnet. Bezogen auf die Frage nach der Genauigkeit der Einschätzung zeigte sich bei den Laien (Tabelle 8), dass die Skala Führungsverhalten im Gegensatz zur Attraktivität der Zielpersonen ($\beta = .64$, $p < .01$) keinen signifikanten Einfluss auf das Intuitivurteil besaß ($\beta = -.13$, n.s.). Auch vom Geschlecht der Zielpersonen ($\beta = -.19$, n.s.) war kein Einfluss auf die intuitive Zuschreibung von Persönlichkeitsstärke zu verzeichnen. Der Interaktionsterm änderte den Aufklärungswert des Gesamtmodells nicht (Änderung in $F = .32$, Änderung in $R^2 = .01$, n.s.). In beiden Schritten betrug die aufgeklärte Varianz 46%.

Bezogen auf das Intuitivurteil der Experten (Tabelle 9) zeigten sich nur leichte Abweichungen. Ebenso wie bei den Laien war die Attraktivität der Zielpersonen eine signifikante Einflussquelle ($\beta = .67$, $p < .01$). Das Geschlecht der Zielpersonen besaß mit einem β von $-.21$ einen marginal signifikanten Einfluss ($p < .10$) auf das Intuitivurteil,

die Skala Führungsstärke dagegen keinen ($\beta = .13$, n.s.). Durch die Überprüfung der Interaktion zwischen Geschlecht und Attraktivität als dritte unabhängige Variable war keine Änderung an der aufklärten Varianz von 50% zu verzeichnen (Änderung in $F = .29$, Änderung in $R^2 = .00$, n.s.).

Variable	Modell 1	Modell 2
Führungsverhalten	.13	.13
Geschlecht	-.19	-.19
Attraktivität	.64**	.64**
Interaktionsterm		-.07
R^2	.46**	.46**

Tabelle 8. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Führungsverhalten, Geschlecht und Attraktivität sowie des Interaktionsterms auf das Intuitivurteil der Laien (*: $p < .05$; **: $p < .01$).

Variable	Modell 1	Modell 2
Führungsverhalten	.13	.13
Geschlecht	-.21	-.21
Attraktivität	.67**	.67**
Interaktionsterm		-.07
R^2	.50***	.50***

Tabelle 9. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Führungsverhalten, Geschlecht und Attraktivität sowie des Interaktionsterms auf das Intuitivurteil der Experten (*: $p < .05$; **: $p < .01$).

Festzuhalten bleibt, dass auf multivariater Ebene die Führungsstärke weder von Experten noch von Laien korrekt erkannt werden konnte.

So wundert es nicht, dass sich die Einschätzungsgüte der Experten, die in ihrem Unternehmen Feedback über die Leistung der von ihnen ausgewählten Mitarbeiter erhielten, nicht von denjenigen Personen unterschied, die kein Feedback erhielten ($t(40) = .35$, n.s.)

Beide Gruppen nutzten die Attraktivität der Zielpersonen als Hinweis auf Führungsstärke. Die Experten orientierten sich darüber hinaus an dem Geschlecht der Zielpersonen, die Laien dagegen nicht. Im letzten Schritt der Datenanalyse stand die Frage im Mittelpunkt, ob Attraktivität und Geschlecht valide Hinweise auf das gezeigte Führungsverhalten darstellten.

Es wurde eine schrittweise Regression mit den unabhängigen Variablen Geschlecht und Attraktivität und der abhängigen Variablen Führungsverhalten berechnet. Weder Geschlecht ($\beta = -.14$, n.s.) noch Attraktivität ($\beta = .23$, n.s.) hatten Auswirkungen auf das tatsächlich gezeigte Führungsverhalten. Weder hatten insgesamt attraktivere Personen noch insgesamt die Männer mehr Führungsverhalten gezeigt. Wie Tabelle 10 zeigt, änderten sich durch die Hinzunahme der Interaktion zwischen Attraktivität und Geschlecht weder das Gewicht der einzelnen Einflussfaktoren noch die aufgeklärte Varianz von insgesamt 6% (Änderung in $F = .02$, Änderung in $R^2 = .00$, n.s.).

Variable	Modell 1	Modell 2
Geschlecht	-.14	-.14
Attraktivität	.23	.23
Interaktionsterm		-.03
R^2	.06	.06

Tabelle 10. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie deren Interaktion auf die tatsächliche Führungsstärke (*: $p < .05$; **: $p < .01$).

Somit konnte das „Körnchen Wahrheit“ in den Stereotypen über Geschlecht und Attraktivität in dieser Studie nicht gefunden werden.

6.5.3. Diskussion

In Studie 3 wurde anhand der Analyse der Beurteilungsgüte von Laien im Vergleich zu Experten untersucht, welche Rolle Expertise in der Personenwahrnehmung spielt. Zu diesem Zweck wurden zwei verschiedene Beurteilerstichproben rekrutiert. Zum einen handelte es sich um eine studentische Stichprobe (Laien) und zum anderen um eine Gruppe erfahrener Personenbeurteiler (Experten). Nur Personen, die seit mindestens einem Jahr beruflich bedingt personelle Auswahlentscheidungen trafen, wurden in diese Gruppe aufgenommen. Nicht nur konnte mit Hilfe dieses Versuchsaufbaus die Genauigkeit der Personenwahrnehmung, sondern darüber hinaus die Stärke und die Richtung der Nutzung des Attraktivitäts- und des Geschlechtsstereotyps im direkten Vergleich zwischen den beiden Gruppen untersucht werden.

Zur Rolle von Expertise bei der Genauigkeit von Persönlichkeitseinschätzungen waren vorab konkurrierende Hypothesen aufgestellt worden. Sollte die Hypothese des Persönlichkeitseinschätzungsinstinktes nach Haselton und Funder (2006) zutreffen, dann müssten beide Gruppen die Führungsstärke der Zielpersonen besser als zufällig einschätzen können, die Gruppen sollten sich jedoch in ihrem Grad der Einschätzungsgenauigkeit nicht unterscheiden. Sollte sich dagegen die Lernperspektive der Intuition (Hogarth, 2001, 2005; Klein, 2004; Plessner et al., 2007) als zutreffend erweisen, so sollten die Experten die Zielpersonen besser hinsichtlich ihrer Führungsstärke einschätzen können.

Als ein erstes wichtiges Ergebnis ist festzuhalten, dass die Experten die Führungsstärke der Zielpersonen nicht besser einschätzen konnten als die Laien. Genauer gesagt konnten weder Laien noch Experten die Führungsstärke der Zielpersonen akkurat erkennen. Auf bivariater Ebene fanden sich zwischen der Laienbeurteilung und der

Führungsstärke der Zielpersonen genauso wie zwischen der Expertenbeurteilung und der Skala Führungsverhalten moderate Zusammenhänge. Doch diese Korrespondenz von .28 (Laien) beziehungsweise .29 (Experten) wurde auf multivariater Ebene mit einem Beta-Koeffizienten von jeweils .13 sowohl bei den Experten als auch bei den Laien noch geringer. Damit erklärt die tatsächliche Führungsstärke der Zielpersonen auf multivariater Ebene lediglich etwa 1% des Antwortverhaltens der Intuitivbeurteiler.

Diese Ergebnisse sprechen somit gegen die Hypothese des Persönlichkeitseinschätzungs-Instinktes, zumindest in Bezug auf die Einschätzung von Führungsstärke.

Selbst Experten, die in ihrem Unternehmen regelmäßig Feedback darüber erhielten, wie gut die von ihnen ausgewählten Bewerber ihre Aufgaben erfüllen, konnten die Führungsstärke der Zielpersonen nicht zutreffend erkennen. Obwohl diese Experten geübt waren im Umgang mit Auswahlentscheidungen und außerdem institutionalisierte Rückmeldungen über die Güte ihrer Entscheidungen erhielten, gelang ihnen keine zutreffende Einschätzung der Zielpersonen. Dieses Ergebnis wiederum widerspricht der Lernperspektive der Intuition, nach der das akkumulierte Erfahrungswissen der Experten sie zu angemesseneren Einschätzungen befähigen sollte.

Ein Grund, warum der auf bivariater Ebene gefundene Genauigkeitseffekt auf multivariater Ebene noch weiter abfiel, lag in dem Einfluss des Attraktivitätsstereotyps begründet. Bei Laien wurde durch die Attraktivität der Zielpersonen alleine etwa 41% des Antwortverhaltens erklärt, bei den Experten etwa 45%. Experten verließen sich somit mitnichten in geringerem Ausmaß auf ihre Stereotype.

Entgegen der Ergebnisse der vorherigen Studien verwendeten die Beurteiler in dieser Studie das Geschlecht der Zielpersonen nicht als Indikator für Führungsstärke. Bei den studentischen Beurteilern hatte das Geschlecht der Zielpersonen keinen Einfluss auf das Urteil, bei den Experten einen marginal signifikanten.

Da die Studie im Rahmen der regulären Arbeitszeit der Experten durchgeführt wurde, bestand die Notwendigkeit, einen eng umgrenzten Zeitrahmen einzuhalten. Aus diesem Grunde wurden 40 der ursprünglich 80 Zielpersonen aus Studie 2 zufällig ausgewählt und als Stimulusmaterial für Studie 3 verwendet. Da Studie 2 auf Basis der 80 Zielpersonen einen signifikanten Einfluss des Geschlechterstereotyps fand, ist davon auszugehen, dass das Ergebnis dieser dritten Studie auf die geringe Stichprobengröße zurückzuführen ist.

Die Ergebnisse dieser Studie finden Rückhalt in anderen Studien, die untersuchten, ob sich erfahrene Beurteiler weniger auf ihre Stereotype verlassen. Marlowe, Schneider und Nelson (1996) fanden beispielsweise, dass erfahrene Personalbeurteiler sich stark an dem Geschlecht und der Attraktivität von Bewerbern orientierten, auch wenn sich in ihrer Studie der Effekt der Attraktivität bei den sehr erfahrenen Personalentscheidern leicht abschwächte.

Attraktiv zu sein zahlt sich somit aus, zumindest in Bezug auf die zugeschriebene Führungsstärke. Dies galt in Studie 3 sowohl für Männer als auch für Frauen. Eine allgemeine Auswirkung von Attraktivität auf die Zuschreibung von Führungsstärke je nach dem Geschlecht der Zielpersonen konnte nicht gefunden werden. Die Vorhersagen des „Lack-of-fit“-Modells (Heilman, 1983; Heilman & Stopeck, 1985) konnten somit nicht bestätigt werden.

Bezieht man in die Analyse des Zusammenwirkens der Merkmale Geschlecht und Attraktivität der Zielpersonen zusätzlich das Geschlecht der Beurteiler ein, so gelangt man zur Überprüfung des „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007). Entgegen der Ergebnisse der ersten beiden Studien fanden sich in Studie 3 insgesamt drei kurvilineare Zusammenhänge, davon zwei bei Laien und einen bei Experten. In der Gesamtschau der Ergebnisse ließ sich der „sexual attribution bias“ jedoch nicht aufzeigen, da sich in drei der vier gleichgeschlechtlichen Bewertungskombinationen ein linearer Zusammenhang als angemessen erwies.

Einer der kurvilinearen Zusammenhänge fand sich bei den studentischen Beurteilern bei der Einschätzung von weiblichen Zielpersonen durch weibliche Beurteiler. Attraktive weibliche Zielpersonen wurden von weiblichen Beurteilern schlechter bewertet als mittelattraktive weibliche Zielpersonen. Dieses Ergebnis entspricht den Annahmen des „sexual attribution bias“.

Die beiden anderen kurvilinearen Zusammenhänge fanden sich jedoch bei der Bewertung von weiblichen Zielpersonen durch männliche Beurteiler und zwar sowohl bei studentischen als auch bei professionellen Beurteilern. Dieses Ergebnis ist nur mit äußerster Vorsicht zu interpretieren, da die Stichprobe der weiblichen Zielpersonen insgesamt gering war ($N = 20$) und zudem die Varianz (.97) und die Standardabweichung (.98) der Attraktivität der weiblichen Zielpersonen in diesem Stimulusmaterial gering waren.

Wenn von den männlichen (studentischen und professionellen) Beurteilern bei attraktiven weiblichen Zielpersonen weniger Führungsstärke wahrgenommen wurde als bei durchschnittlich attraktiven Zielpersonen, so könnte dies vorsichtig als modifizierte Version des „Lack-of-fit“-Modells (Heilman, 1983) interpretiert werden.

So verzeichnen einige Forschungsergebnisse auf Seiten der Frauen einen Wandel des zugrunde liegenden „think manager – think male“-Stereotyps (Diekmann et al., 2005; Schein et al., 2001; Sczesny, 2003a). Sollten somit hauptsächlich Männer noch die Rolle einer Führungskraft und Führungsstärke als typisch maskulin ansehen, dann könnte es aus der männlichen Sichtweise gerade bei attraktiven und damit gleichzeitig femininen (Gillen, 1981) Frauen zu einer Wahrnehmung von Diskrepanz kommen. So wäre es prinzipiell denkbar, dass sich das „Lack-of-fit“-Modell weiterhin als zutreffend erweist, jedoch lediglich für männliche Beurteiler.

Nachdem insgesamt festgestellt wurde, dass beide Gruppen von Beurteilern die Führungsstärke der Zielpersonen intuitiv nicht erkennen konnten, sondern der „what is beautiful is good“-Heuristik

und zumindest im Ansatz dem „think manager - think male“- Stereotyp folgten, wurde schließlich untersucht, ob in diesen Daumenregeln ein „Körnchen Wahrheit“ steckte und somit die Verwendung dieser Indikatoren als sinnvolle Strategie anzusehen wäre. Allerdings war auch dies nicht der Fall. Attraktivere Zielpersonen zeichneten sich nicht durch mehr führungsstarke Verhaltensweisen innerhalb der Verhaltensmessung aus. Genauso hatte das Geschlecht keinen Einfluss auf das tatsächliche Führungsverhalten.

Resultierend aus den bisherigen Ergebnisse dieser Forschungsreihe wäre in Bezug auf die vorliegende Fragestellung uneingeschränkt der sozialen Kognitionsforschung zuzustimmen. Menschen verlassen sich bei der intuitiven Einschätzung der Führungsstärke fremder Personen vornehmlich auf ihre Stereotype. Darüber hinaus erkennen sie die wahre Führungsstärke jedoch nicht. Auch führte die Nutzung der Stereotype als simple und sparsame Heuristik nicht zu guten Einschätzungen, da ihnen das „Körnchen Wahrheit“ fehlt. Die Stereotype über Geschlecht und Attraktivität fanden keine Entsprechung in der Realität. Zumindest nicht mit der in Studie 2 und 3 gemessenen Realität.

Ein plausibler Einwand gegen die bisher gefundenen Ergebnisse könnte lauten, dass Menschen flexibel in der Übernahme von Rollen seien. So könnte jemand, der in der Turmbauübung bei dieser speziellen Aufgabe und in dieser speziellen Gruppe nur eine folgende Rolle eingenommen hat, in einer anderen Situation sehr wohl eine Führungsrolle einnehmen. So könnte es ein Fehlschluss sein, jemandem Führungsfähigkeiten abzusprechen, der in der Turmbauübung wenig Führungsverhalten gezeigt hat (Heilmann, 2002).

Weiterhin könnte man argumentieren, dass Attraktivität und Geschlecht zwar in diesen Studien keine Auswirkungen auf die tatsächliche Führungsstärke hatten, dies aber in der Realität außerhalb des Labors und außerhalb einer Gruppenübung mit Studenten sehr wohl der Fall sein könnte. Beispielsweise zeigen verschiedene Studien,

dass kontextuelle Faktoren einen Einfluss darauf haben, ob Männer oder Frauen in einer Gruppe eher die Führungsrolle übernehmen. Bei sehr strukturierten oder maskulinen Aufgaben (z.B. das Reparieren einer Maschine) übernehmen Männer eher die Führungsrolle, bei eher femininen Aufgaben (z.B. künstlerische) oder Aufgaben, die komplexe soziale Interaktionen (z.B. Verhandeln, Teilen und Koordinieren von Ideen) beinhalten, übernehmen dagegen eher Frauen die Führungsrolle (Eagly & Karau, 2002). Zusammengefasst zeigen Männer mehr Führungsverhalten bei aufgabenorientierten Gruppenübungen, Frauen zeigen dagegen mehr soziales Führungsverhalten.

Die in diesen Studien realisierte Turmbauübung stellte eine Kombination aus beiden situationalen Anforderungen dar. Die prinzipiell gruppendynamisch orientierte Turmbauübung wurde durch die Konkurrenz mit anderen Gruppen um den höchsten, stabilsten und kreativsten Turm mit einer zusätzlichen Wettbewerbsorientierung versehen, die ein gewisses Maß an Aufgabenorientierung notwendig machte. So ist es prinzipiell denkbar, dass nur aufgrund dieser speziellen Gruppenaufgabe Frauen und Männer sich nicht anhand ihres Führungsverhaltens unterschieden. Studie 4 diente daher dem Ziel, diese Ergebnisse in eine weitere Realität zu übertragen.

6.6. Studie 4

Mit Hilfe von Studie 4 sollten die bisherigen Ergebnisse um zwei wesentliche Aspekte erweitert werden. Zum einen wurden bisher als Zielpersonen ausschließlich Studenten eingeschätzt. Da sich Studenten noch in ihrer Ausbildung befinden und daher in der Regel noch keine Daten über ihre tatsächliche Führungsstärke vorliegen, wurden für eine möglichst valide Messung zweierlei Maßnahmen angesetzt.

Erstens wurde auf Seiten der intuitiven Beurteiler als Einschätzungsinstrument eine Skala gewählt, die Führungsstärke als Disposition unabhängig von Alter, Erfahrung und Soziodemographie (Noelle-

Neumann, 1983) erfasst und daher auch für die Einschätzung von Studenten angemessen war.

Zweitens wurde auf Seiten der Zielpersonen konkretes Führungsverhalten durch eine Assessment Center Übung induziert, welches mit Hilfe der Skala Führungsverhalten dokumentiert wurde und somit als Referenzmaß für wahre Führungsstärke zur Verfügung stand.

Um den Geltungsbereich der bisherigen Ergebnisse zu erweitern, fungierten in Studie 4 erwachsene – im Berufsleben stehende – Menschen als Zielpersonen. Diese strukturell unterschiedliche Zielpersonengruppe bot eine weitere Möglichkeit, die gleichsam den zweiten wesentlichen Erweiterungsaspekt der bisherigen Ergebnisse beinhaltet. Da diese Personen bereits im Berufsleben stehen, ist ihr tatsächliches Gruppen- und Führungsverhalten täglich beobachtbar. Diese Tatsache wurde sich zunutze gemacht, um ein Referenzmaß für Führungsstärke zu generieren, welches auf dem überdauernden tatsächlichen Verhalten der Zielpersonen in ihrem Beruf beruhte.

6.6.1. Ein Referenzmaß für Führungsstärke II: das 360°-Feedback

Durch den Gang in das Arbeitsleben von Menschen und der dortigen Rekrutierung von Zielpersonen eröffnete sich ein weiteres Feld an möglichen Referenzkriterien für Führungsstärke. Wichtigstes Kriterium hierbei war eine möglichst umfassende Abbildung der Leistung und des Verhaltens der Arbeitnehmer.

In diesem Rahmen hat sich in den letzten Jahren in der Feedbackkultur vieler Unternehmen das so genannte 360°-Feedback eingebürgert (Fleenor, Taylor & Chappelow, 2008). Der Name 360°-Feedback soll die vielen Richtungen widerspiegeln, aus denen Rückmeldungen auf Basis von vorher definierten Evaluationsdimensionen stammen (Balyan, Kanwar & Pant, 2005). Das Prinzip des 360°-Feedbacks besteht darin, dass anhand eines standardisierten Fragebogens regelmäßig Einschätzungen von verschiedenen Gruppen über eine

Person eingeholt werden. Zu diesen Gruppen gehören in der Regel die Linienvorgesetzten, Kollegen und - falls vorhanden - die direkten Untergebenen und die internen und/oder externen Kunden (Scherer & Kaufel, 2005).

Im Mittelpunkt einer 360°- Beurteilung steht im Allgemeinen die berufliche Kompetenz der Fokusperson, das heißt ihr kriterien- und leistungsbezogenes Verhalten im Sinne von Fähigkeiten und Fertigkeiten (Scherer & Kaufel, 2005). Dieses multiperspektivische Feedback dient im unternehmerischen Alltag vor allem zur Bestimmung von Stärken und Schwächen, Entwicklungsbedarf und Entwicklungsmöglichkeiten, Leistungs-, Führungs- und Personalbeurteilungen (Rao & Rao, 2005).

Ansätze zur Integration mehrerer Feedbackquellen für eine umfassende Abbildung der Leistung eines Mitarbeiters gibt es schon sehr lange. Der Vorteil des 360°-Feedbacks besteht in der systematischen Integration aller verfügbaren Quellen, um die jeweiligen Mängel der Einzelquellen auszugleichen (Marcus & Schuler, 2006).

Vorgesetzte beispielsweise wissen zwar um die allgemeine Leistung ihrer Mitarbeiter, dennoch haben sie meist nicht die Zeit, um ihre Untergebenen intensiv und differenziert zu beobachten. Nicht immer haben Vorgesetzte einen vollständigen Einblick in die Tätigkeit ihrer Mitarbeiter (Rao & Rao, 2005). Dies gilt insbesondere für Spezialistentätigkeiten. Hinzu kommt, dass Untergebene meist viel daran setzen, vor ihrem Vorgesetzten gut auszusehen, was wiederum abträglich für eine realistische Vorgesetztenbeurteilung ist.

Die Befragung der Kollegen dagegen hat den Vorteil, dass diese eher die Möglichkeit haben, die Fokusperson bei der täglichen Arbeit zu beobachten und Arbeitsabläufe zu bewerten. Allerdings haben auch Kollegen nicht immer einen vollständigen Einblick in die Tätigkeit und weiterhin können deren Urteile durch Sympathie oder Antipathie sowie durch die eigenen arbeitsrelevanten Belange verzerrt sein (Marcus & Schuler, 2006).

Für das in dieser Studie als Referenzkriterium verwendete 360°-Feedback wurden daher sowohl die Vorgesetzten als auch die Kollegen der Zielpersonen befragt. Handelte es sich bei der Fokusperson um eine Führungskraft, so wurden weiterhin deren Mitarbeiter befragt. Gerade Führungskräfte haben oftmals nur ein unvollständiges Bild ihrer eigenen Leistung, da sie seltener eine qualifizierte und objektive Rückmeldung bekommen. Allerdings sind die Aussagen von Untergebenen oftmals beschönigt, da sich Mitarbeiter nie vollkommen sicher sein können, was innerhalb des Unternehmens mit ihren Aussagen geschieht (Marcus & Schuler, 2006).

In Studie 4 wurde daher ein 360°-Feedback zur Objektivierung von Führungsstärke realisiert, welches die Perspektiven der Linienvorgesetzten, der Kollegen und – falls vorhanden – der Mitarbeiter einschloss. Diese Herangehensweise hat verschiedene Vorteile.

Erstens ist ein 360°-Feedback wesentlich vollständiger und objektiver als die Einschätzung einer Einzelperson oder einer einzelnen Gruppe, da sich die Grenzen der einzelnen Perspektiven gegenseitig ausgleichen (Fleenor et al., 2008). Zweitens verhalten sich Menschen nicht immer gleich, sondern entsprechend der Rolle, die sie gerade einnehmen. Dieses unterschiedliche Verhalten, die jeweiligen Stärken und Schwächen einer Person in der jeweiligen Rolle können durch ein 360°-Feedback abgebildet werden (Rao & Rao, 2005). Beispielsweise kann eine Führungskraft im Umgang mit Kollegen der gleichen Führungsebene sehr geschickt sein, während ihr Einfühlungsvermögen bei ihren Mitarbeitern wenig ausgeprägt ist. Drittens bleibt im 360°-Feedback durch die Aggregation der Einzelwerte zu einem Gruppenmittelwert oder sogar zu einem Gesamtmittelwert über alle Gruppen der einzelne Bewerter anonym. Insofern ist anzunehmen, dass die Einschätzung offener und ehrlicher erfolgt (Scherin & Kaufel, 2005).

Zusammenfassend wurden für die vierte Studie in verschiedenen Unternehmen Mitarbeiter und Führungskräfte als Zielpersonen auf Video aufgenommen. Anschließend wurde ein 360°-Feedback über sie

erhoben, indem jeweils eine gewisse Anzahl ihrer jeweiligen Linien-vorgesetzten, Kollegen und - falls vorhanden - Untergebenen um eine Einschätzung gebeten wurden. Um die jeweiligen Verzerrungen der einzelnen Perspektiven zu nivellieren, wurden die Einschätzungen der einzelnen Gruppen gemittelt.

Zusätzlich zu dieser multiperspektivischen Betrachtung als Referenzmaß für Führungsstärke wurde die tatsächliche berufliche Position der Zielpersonen erhoben. Unterschieden wurden hierbei *Mitarbeiter* und *Führungskräfte*, welche als Personen mit disziplinarischer Führungsverantwortung definiert wurden. Als Führungskräfte galten somit Zielpersonen denen, unabhängig von der jeweiligen Führungsebene, in direkter Linie Mitarbeiter unterstellt waren.

6.6.2. Methode

Zielpersonen

Im ersten Schritt der Datengenerierung wurden die Zielpersonen gefilmt. Diese waren Arbeitnehmer aus drei verschiedenen Unternehmen innerhalb Deutschlands. Die Unternehmen entstammten den drei Branchen Versicherung, Marktforschung und Telekommunikation mit Sitz in drei verschiedenen Städten, namentlich Köln, Hamburg und München. Als Stimulusmaterial wurden insgesamt 81 Personen im Alter zwischen 22 und 60 Jahren, im Durchschnitt 39.70 Jahre ($SD = 3.23$), gefilmt, darunter 34 Frauen und 47 Männer. Unter diesen befanden sich 45 Mitarbeiter ohne Führungsverantwortung (im Folgenden *Mitarbeiter* genannt) und 36 Mitarbeiter mit disziplinarischer Führungsverantwortung (im Folgenden *Führungskräfte* genannt). Die 24 weiblichen und 21 männlichen Mitarbeiter waren zwischen 22 und 60 Jahre alt, im Durchschnitt 38.29 ($SD = 9.22$). Die Gruppe der Führungskräfte bestand aus 10 Frauen und 26 Männern im Alter zwischen 30 und 55 Jahren, im Durchschnitt 41.47 ($SD = 6.49$).

Erstellt wurde das Videomaterial in dem jeweiligen Unternehmen vor Ort. Hierzu fanden sich die Zielpersonen einzeln jeweils in einem von dem Unternehmen zur Verfügung gestellten Raum ein. Um zu verhindern, dass Führungskräfte an ihrem Kleidungsstil erkannt werden, trugen alle Zielpersonen einheitliche Sweatshirtjacken. Bei den Videoaufnahmen saßen die Zielpersonen an einem Tisch vor einer weißen Wand, so dass auf dem Filmausschnitt Kopf, Oberkörper und Arme zu sehen waren. Mit Blick in die Kamera sollten sie etwa eine Minute lang ihren beruflichen Werdegang beschreiben, welche Tätigkeit sie aktuell im Unternehmen ausführen und was sie in ihrer Freizeit unternehmen. Die Anonymität der Zielpersonen wurde dabei durch zwei Maßnahmen gewährleistet. Erstens trugen sie einen Button mit einer Versuchspersonennummer an ihrer Jacke, so dass alle erhobenen Daten nur dieser Nummer zugeordnet werden konnten. Zweitens war es den Zielpersonen freigestellt, bei der Vorstellung in die Kamera ihren Namen zu nennen.

Zur Fertigstellung des Stimulusmaterials wurde das Videomaterial der 81 Personen auf jeweils 20-sekündige tonlose Sequenzen gekürzt und zu einem Gesamtfilm zusammengefügt. Dabei wurde der Bildschirm nach jeder Zielperson für die Dauer von 9 Sekunden schwarz. Die nächste Zielperson wurde mit der laufenden Nummer angekündigt. Die Gesamtdauer des Films betrug etwa 40 Minuten.

Die Attraktivität der Zielpersonen wurde anschließend auf Basis dieses Films durch 9 weibliche und 6 männliche Studenten der Universität zu Köln im Alter zwischen 22 und 30 Jahren, im Durchschnitt 25.40 Jahre ($SD = 1.88$), bewertet. Unter den 15 Beurteilern führte ein Intraklassen-Koeffizient von .59 insgesamt zu einem Cronbach's Alpha von .95, so dass die Einzelwerte zu einem Mittelwert pro Zielperson aggregiert wurden. Weibliche Zielpersonen wurden auf die Frage „Person 1-81 finde ich...“ auf einer siebenstufigen Skala von (1) „überhaupt nicht attraktiv“ bis (7) „sehr attraktiv“ (Anhang C) mit einem Mittelwert von 3.61 ($SD = .98$) bewertet, männliche Zielperso-

nen mit $M = 3.04$ ($SD = .82$). Im Durchschnitt wurde den Zielpersonen eine Attraktivität in Höhe von 3.27 ($SD = .93$) attestiert.

Das 360°-Feedback

Im Anschluss an die Videoaufnahmen wurden jeder Zielperson fünf Umschläge mitgegeben mit der Bitte, die enthaltenen Fragebögen von ihren Kollegen, direkten Vorgesetzten und - falls vorhanden - Mitarbeitern ausfüllen zu lassen. Auf diesen fünf Fragebögen war jeweils die Versuchspersonennummer entsprechend des im Video getragenen Buttons notiert. Um ein offenes Antwortverhalten des Arbeitsumfeldes der jeweiligen Zielperson zu gewährleisten, wurden die Fragebögen nicht an die Zielperson selber beziehungsweise an die Personalabteilung zurückgegeben, sondern direkt an die Versuchsleiter. Diese Maßnahme diente der Sicherstellung eines möglichst umfassenden und realistischen Fremdbildes. Die Fragebögen zur Erhebung des 360°-Feedbacks enthielten eine Abfrage demographischer Daten der ausfüllenden Person, ein Item zur Erfassung der beruflichen Beziehung zwischen der einschätzenden Person und der Zielperson, sowie die drei Items der Skala Persönlichkeitsstärke nach Noelle-Neumann (Anhang E).

Beurteiler

Die Persönlichkeitsstärke der Zielpersonen wurde an der Universität zu Köln von 26 weiblichen und 22 männlichen Studenten im Alter von 21 bis 33, durchschnittlich 24.88 ($SD = 2.44$), Jahren eingeschätzt. Weitere 52 Beurteiler, davon 29 weiblich und 23 männlich, im Alter von 21 bis 29 Jahren ($M = 24.33$, $SD = 1.64$) schätzten die berufliche Position der Zielpersonen ein.

Versuchsaufbau

Wie bei allen vorhergehenden Untersuchungen wurde den Intuitivbeurteilern die Information an die Hand gegeben, dass es sich bei dem Versuch um eine Studie zum ersten Eindruck handele, sie die Zielpersonen rein auf Basis dieses Eindrucks einschätzen sollten und es

somit keine richtigen oder falschen Antworten gäbe. Auf dem hierfür vorgelegten Fragebogen befand sich die Instruktion noch einmal in schriftlicher Form. Nach jedem Videoclip beantworteten die Beurteiler intuitiv die für ihre Dimension vorgesehenen Fragen in Bezug auf die soeben gesehene Zielperson.

Die Führungsstärke der Zielpersonen wurde analog zu Studie 1-3 und zum 360°-Feedback mit den drei Items der Skala Persönlichkeitsstärke nach Noelle-Neumann erhoben (Anhang B).

Der vermutete berufliche Status der Zielpersonen wurde mit dem Item „Die Person 1-81 ist eine Führungskraft“ auf einer siebenstufigen Skala mit den Polen (1) „Trifft überhaupt nicht zu“ bis (7) „Trifft voll und ganz zu“ abgefragt.

Nach der Einschätzung aller 81 Zielpersonen wurde den Beurteilern für ihre Versuchsteilnahme gedankt und sie wurden über das Ziel der Studie informiert.

6.6.3. Ergebnisse

Zunächst wurde auf allen zuvor beschriebenen Dimensionen die Einigkeit der Einschätzungen überprüft.

Beginnend mit dem 360°-Feedback über die Zielpersonen zeigte sich, dass sich die einzelnen Gruppen im Hinblick auf die Persönlichkeitsstärke der Zielpersonen wenig einig waren. Bezogen auf die Führungskräfte innerhalb dieser Stichprobe ergab sich für die Einschätzung durch die Vorgesetzten, Mitarbeiter und Kollegen ein Intraklassen-Koeffizient von .21, welcher zu einem Cronbach's Alpha von .44 führte. Dieser geringe Wert ist vermutlich vor allem auf den geringen Zusammenhang zwischen Führungskräfte- und Mitarbeiter-einschätzungen ($r = .21$, n.s.) zurückzuführen, der geringer ausgeprägt war als der Zusammenhang der Einschätzungen von Kollegen und Mitarbeitern ($r = .37$, n.s.) sowie zwischen den Einschätzungen durch Kollegen und Führungskräfte ($r = .37$, $p < .05$).

In Bezug auf die Mitarbeiter resultierte die Einschätzung durch Vorgesetzte und Kollegen in einen Intraklassen-Koeffizient von .42, der zu einem Cronbach's Alpha von .59 führte.

Diese geringe Übereinstimmung der verschiedenen Perspektiven im 360°-Feedback unterstreicht die Notwendigkeit einer Mittelung der verschiedenen Feedbackquellen, um ein objektives Maß der Führungsstärke der Zielpersonen zu generieren. Für die folgenden Berechnungen wurden daher die unterschiedlichen Perspektiven zu einem Wert pro Zielperson aggregiert.

Im Hinblick auf die beiden Dimensionen Persönlichkeitsstärke und berufliche Position ließ sich unter den Intuitivbeurteilern eine höhere Einigkeit verzeichnen. Bei der Einstufung der beruflichen Position der Zielpersonen auf der siebenstufigen Skala ergab sich ein Intraklassen-Koeffizient von .31, der zu einem Cronbach's Alpha von .96 führte. Bei der Zuschreibung von Persönlichkeitsstärke fand sich unter den 48 Beurteilern ein Intraklassen-Koeffizient von .27 (Cronbach's Alpha = .95). Da somit die Interraterreliabilität jeweils als ausreichend anzusehen war, wurden für jedes eingeschätzte Merkmal die Einzelwerte zu einem Mittelwert pro Zielperson aggregiert.

Bei der Identifikation möglicher Orientierungspunkte für die Zuschreibung von Persönlichkeitsstärke wurden zunächst die potenziellen Einflüsse des Attraktivitäts- und des Geschlechtsstereotyps analysiert.

Dabei zeigte sich, dass die intuitiven Urteile an der Attraktivität der Zielpersonen orientiert waren. Zwischen der Attraktivität der Zielpersonen und der zugeschriebenen Persönlichkeitsstärke ergab sich ein signifikanter Zusammenhang von $r = .43$ ($p < .001$). Dagegen ergab sich kein signifikanter Zusammenhang zwischen dem Geschlecht der Zielpersonen und der intuitiv zugeschriebenen Persönlichkeitsstärke ($r = -.12$, n.s.). Weiterhin zeigte sich durch eine signifikant positive Relation zwischen Geschlecht und Attraktivität der Zielpersonen ($r = .30$, $p < .01$) ein Zusammenwirken dieser beiden Variablen.

Für die weitere Untersuchung des Wahrnehmungsprozesses wurde auf multivariater Ebene der Einfluss der beiden Variablen auf die spontane Einschätzung von Persönlichkeitsstärke geprüft. Zu diesem Zweck wurden im ersten Schritt der Regression mit der abhängigen Variablen Intuitivurteil die Einflüsse der beiden unabhängigen Variablen Geschlecht und Attraktivität unabhängig voneinander untersucht. Durch das Verschwinden des Suppressor-Effekts in der multivariaten Analyse zeigte sich, dass beide Merkmale mit einem β von .47 ($p < .001$) für Attraktivität und $-.26$ ($p < .05$) für Geschlecht jeweils einen signifikanten Einfluss auf das Intuitivurteil besaßen.

Zur Analyse der Annahmen des „Lack-of-fit“-Modells zum Zusammenwirken der beiden Personenmerkmale Attraktivität und Geschlecht wurde zunächst ein Interaktionsterm gebildet, indem Attraktivität und Geschlecht erst zentriert und dann miteinander multipliziert wurden. Dieser wurde im zweiten Schritt der Regression als dritte unabhängige Variable hinzugenommen. Die Veränderungsparameter des Modells weisen darauf hin, dass das Zusammenwirken der beiden Stereotype keinen Erklärungswert für das Gesamtmodell besaß ($\beta = .02$, Änderung in $F = .03$, Änderung in $R^2 = .00$, n.s.).

Darüber hinaus wurde untersucht, ob ein „sexual attribution bias“ vorlag. Dazu wurden zunächst getrennt nach Geschlecht der Beurteiler und der Zielpersonen bivariate Korrelationen zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke berechnet.

Es zeigte sich, dass die Zusammenhänge in allen vier möglichen Kombinationen positiv und signifikant waren. Bei der Beurteilung von weiblichen Zielpersonen wurde die Attraktivität in gleichem Ausmaß von männlichen Beurteilern ($r = .60$, $p < .001$) wie von weiblichen Beurteilern ($r = .53$, $p < .001$) genutzt ($Z = .33$, n.s.). Bei männlichen Zielpersonen wurde der Attraktivität bei der Einschätzung durch gegengeschlechtliche Beurteiler eine ebenso hohe Relevanz ($Z = .16$, n.s.) beigemessen ($r = .45$, $p < .01$) wie bei der Bewertung durch gleichgeschlechtliche Beurteiler ($r = .41$, $p < .01$).

Da der „sexual attribution bias“ vorhersagen würde, dass der Zusammenhang zwischen Attraktivität und intuitiv wahrgenommener Persönlichkeitsstärke bei der gegengeschlechtlichen Beurteilung linear, bei der gleichgeschlechtlichen dagegen kurvilinear ausfallen sollte, wurden Kurvenanpassungen für die vier möglichen Einschätzungskombinationen (1) weibliche Zielpersonen und weibliche Beurteiler, (2) weibliche Zielpersonen und männliche Beurteiler, (3) männliche Zielpersonen und weibliche Beurteiler und (4) männliche Zielpersonen und männliche Beurteiler berechnet.

Wie Tabelle 11 zeigt, führte die Modellierung eines kurvilinearen Zusammenhangs in keiner der vier Kombinationen zu einer signifikant besseren Modellanpassung. Die Voraussagen des „sexual attribution bias“ konnten somit nicht bestätigt werden.

Zusammenhang	<i>weibliche Beurteiler, weibliche Zielpersonen</i>	<i>Weibliche Beurteiler, männliche Zielpersonen</i>	<i>Männliche Beurteiler, weibliche Zielpersonen</i>	<i>Männliche Beurteiler, männliche Zielpersonen</i>
Linear	28%	20%	36%	17%
Quadratisch	31%	20%	38%	17%
Steigers Z	-1.13	-.05	-.84	-.26

Tabelle 11. Aufgeklärte Varianz bei Annahme eines linearen versus quadratischen Zusammenhangs zwischen Attraktivität und zugeschriebener Persönlichkeitsstärke getrennt nach Geschlecht der Beurteiler und Geschlecht der Zielperson (*: $p < .05$; **: $p < .01$).

Somit lässt sich als Zwischenfazit festhalten, dass die Attraktivität und das Geschlecht für die Beurteiler als Hinweise auf die Führungsstärke der Zielpersonen fungierten und dies unabhängig von dem Geschlecht der Zielpersonen und dem Geschlecht der Beurteiler. Beide Merkmale übten dabei unabhängig voneinander einen direkten

Einfluss auf die zugeschriebene Führungsstärke aus. Damit wurden die Voraussagen des „think manager – think male“ und des „what is beautiful is good“- Stereotyps bestätigt, die des „Lack-of-fit“- Modells und des „sexual attribution bias“ dagegen nicht.

Die Art der generierten Daten in dieser Studie ermöglichten über die Analyse des Beurteilungsprozesses hinaus zwei unterschiedliche Herangehensweisen zur Ermittlung der Genauigkeit der Personenwahrnehmung. Zunächst wurde die Übereinstimmung der Intuitivbeurteilung mit dem 360°-Feedback als Referenzkriterium für Führungsstärke betrachtet.

Zunächst zeigte sich bei dieser Betrachtung, dass es einen Zusammenhang zwischen beruflicher Position und 360°-Feedback gab. Ein t-Test für unabhängige Stichproben zeigte, dass Führungskräfte im 360°-Feedback einen signifikant höheren Persönlichkeitsstärkewert ($M = 5.40$, $SD = .77$) als Mitarbeiter ($M = 4.81$, $SD = .67$) erreichten ($t(79) = -3.65$, $p < .001$). Dies wurde von den intuitiven Beurteilern in gleicher Weise wahrgenommen. Sie schrieben den Führungskräften mit einem Mittelwert von 4.20 ($SD = .70$) eine signifikant höhere Persönlichkeitsstärke ($t(79) = -3.18$, $p < .01$) zu als Mitarbeitern mit einem Mittelwert von 3.72 ($SD = .65$).

Die bivariate Korrelation zwischen zugeschriebener Persönlichkeitsstärke und den Werten des 360°-Feedbacks ergab einen signifikanten Zusammenhang von $r = .29$ ($p < .01$). Zielpersonen, die intuitiv als persönlichkeitsstark wahrgenommen wurden, bekamen somit von ihrem Arbeitsumfeld ebenfalls eine hohe Persönlichkeitsstärke zugeschrieben.

Für die multivariate Betrachtung wurde eine schrittweise Regression durchgeführt. Im ersten Schritt wurden die unabhängigen Variablen 360°-Feedback, Geschlecht und Attraktivität auf ihren Aufklärungswert bezüglich der abhängigen Variablen zugeschriebene Persönlichkeitsstärke hin untersucht. Die beiden Zielpersonenmerkmale Geschlecht ($\beta = -.26$, $p < .05$) und Attraktivität ($\beta = .47$, $p < .01$) besaßen

beide einen signifikanten Einfluss auf das Intuitivurteil. Darüber hinaus wurde das 360°-Feedback ($\beta = .18$, $p < .08$) als Prädiktor marginal signifikant. Im zweiten Schritt sollte die Aufnahme des Interaktionsterms als dritte unabhängige Variable überprüfen, ob Attraktivität in multiplikativer Verknüpfung mit dem Geschlecht einen zusätzlichen Erklärungswert für die Varianz der Intuitivurteile darstellt. Dies war jedoch nicht der Fall (Änderung in $F = .02$, Änderung in $R^2 = .00$, n.s.).

Variable	Modell 1	Modell 2
360°-Feedback	.18	.18
Geschlecht	-.26*	-.26*
Attraktivität	.47**	.46**
Interaktionsterm		.02
R^2	.28**	.28**

Tabelle 12. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen 360°-Feedback, Geschlecht und Attraktivität sowie des Interaktionsterms auf das Intuitivurteil (*: $p < .05$; **: $p < .01$).

Zur Überprüfung der „Körnchen Wahrheit“- Hypothese wurde im weiteren Verlauf der Datenanalyse berechnet, inwieweit den beiden Merkmalen Attraktivität und Geschlecht sowie deren Interaktion ein Einfluss auf die Werte des 360°-Feedbacks zukam. Eine schrittweise Regression mit den unabhängigen Variablen Geschlecht und Attraktivität und der abhängigen Variablen 360°-Feedback demonstrierte, dass Attraktivität einen signifikanten Einfluss besaß ($\beta = .26$, $p < .05$). Ein signifikanter Einfluss des Geschlechts konnte dagegen nicht gefunden werden ($\beta = -.09$, n.s.). Auch die im zweiten Schritt hinzugenommene Interaktion zwischen beiden Merkmalen als dritte unabhängige Variable hatte keinen signifikanten Erklärungswert (Änderung in $F = 1.41$, Änderung in $R^2 = .02$, n.s.).

Variable	Modell 1	Modell 2
Geschlecht	-.09	-.08
Attraktivität	.26*	.28*
Interaktionsterm		-.13
R^2	.06	.08

Tabelle 13. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie des Interaktionsterms auf das 360°-Feedback (*: $p < .05$; **: $p < .01$).

So kann als Fazit der ersten Überprüfung der Genauigkeit der Einschätzung von Führungsstärke festgehalten werden, dass die Beurteiler sich sowohl am Geschlechts- als auch am Attraktivitätsstereotyp orientierten, darüber hinaus jedoch zu einem gewissen Grad die Führungsstärke erkennen konnten. Auch zeugte die Attraktivität als Indikator von einer gewissen Validität, da attraktivere Personen auch im 360°-Feedback höhere Werte erlangten.

Neben der Überprüfung der Übereinstimmung der Intuitivbeurteilung mit dem 360°-Feedback wurde eingangs eine zweite Möglichkeit zur Betrachtung der Wahrnehmungsgenauigkeit der intuitiven Beurteiler skizziert. Daher wurde im nächsten Schritt die Übereinstimmung der intuitiv vermuteten beruflichen Position mit der tatsächlichen Position der Zielpersonen verglichen.

Wie eingangs erwähnt, hatten die Intuitivbeurteiler die berufliche Position der Zielpersonen auf einer siebenstufigen Skala eingeschätzt. Hinsichtlich der intuitiv vermuteten beruflichen Position wurden nun Mitarbeiter und Führungskräfte mit Hilfe eines t-Tests für unabhängige Stichproben verglichen. Es zeigte sich, dass den Führungskräften mit einem Mittelwert von 3.85 ($SD = .83$) ein signifikant höherer Wert ($t(79) = -3.66$, $p < .001$) zugeschrieben wurde als den Mitarbeitern mit einem Mittelwert von 3.22 ($SD = .73$).

Das intuitive Erkennen der beruflichen Position der Zielpersonen konnte auch auf multivariater Ebene nachgewiesen werden. Im ersten Schritt wurde eine Regression mit der abhängigen Variablen vermutete berufliche Position und der unabhängigen Variablen tatsächliche berufliche Position berechnet. Es zeigte sich, dass die tatsächliche berufliche Position alleine 15% des Antwortverhaltens der intuitiven Beurteiler erklärte und einen signifikanten Einfluss besaß ($\beta = .38$, $p < .01$). Im zweiten Schritt der Regression wurden die beiden unabhängigen Variablen Geschlecht und Attraktivität der Zielpersonen aufgenommen. Hier zeigte sich, dass auch diese beiden Einflussgrößen mit einem β von $-.30$ für Geschlecht und $.36$ für Attraktivität als Prädiktoren signifikant wurden ($p < .01$). Doch auch die tatsächliche berufliche Position der Zielpersonen wurde signifikant ($\beta = .27$, $p < .01$). Die Hinzunahme der beiden unabhängigen Variablen hatte einen signifikanten Einfluss auf das Gesamtmodell (Änderung in $F = 7.61$, Änderung in $R^2 = .14$, $p < .01$) und erhöhte die erklärte Varianz auf 29%.

Die im dritten Schritt aufgenommene Interaktion zwischen Geschlecht und Attraktivität hatte keine signifikante Auswirkung auf das Gesamtmodell (Änderung in $F = 1.86$, Änderung in $R^2 = .01$, n.s.).

Variable	Modell 1	Modell 2	Modell 3
Berufliche Position	.38**	.27**	.28**
Geschlecht		-.30**	-.30**
Attraktivität		.36**	.33**
Interaktionsterm			.13
R^2	.15**	.29**	.30**

Tabelle 14. Beta-Koeffizienten der schrittweisen Regression zur Analyse der Prädiktorvariablen tatsächliche berufliche Position, Geschlecht und Attraktivität sowie deren Interaktion auf die vermutete berufliche Position (*: $p < .05$; **: $p < .01$).

Zur Überprüfung der „Körnchen Wahrheit“- Hypothese der verwendeten Stereotype wurden die Auswirkungen der beiden Merkmale Attraktivität und Geschlecht sowie deren Interaktion auf die tatsächliche berufliche Position untersucht. Eine schrittweise binär logistische Regression mit den unabhängigen Variablen Geschlecht und Attraktivität und der abhängigen Variablen berufliche Position zeigte, dass das Geschlecht in beiden Modellen einen signifikanten Einfluss auf die tatsächliche berufliche Position hatte (Modell 1 Regressionskoeffizient $B = -.69$, $p < .01$; Modell 2 = $-.67$, $p < .05$). Die Wahrscheinlichkeit, eine Führungskraft zu sein, war in der untersuchten Stichprobe als Mann um den Faktor 0.51 höher. Der Einfluss der Attraktivität wurde lediglich in Modell 2 marginal signifikant (Regressionskoeffizient $B = .44$, $p < .10$). Die Interaktion zwischen Geschlecht und Attraktivität hatte keinen signifikanten Einfluss.

Variable	Modell 1	Modell 2
Geschlecht	-.69**	-.67*
Attraktivität	.42	.44
Interaktionsterm		-.15
<i>Nagelkerkes R^2</i>	.13	.13

Tabelle 15. Regressionskoeffizienten der binär logistischen Regression zur Analyse der Prädiktorvariablen Geschlecht und Attraktivität sowie deren Interaktion auf die tatsächliche berufliche Position (*: $p < .05$; **: $p < .01$).

6.6.4. Diskussion

Studie 4 diente der Erweiterung der bisherigen Ergebnisse um zwei wesentliche Aspekte. Zum einen wurde über studentische Zielpersonen hinausgehend eine Stichprobe von Arbeitnehmern rekrutiert. Zum anderen erlaubte diese Maßnahme gleichsam andere Messungen und damit andere Referenzkriterien für Führungsstärke.

Da die bereits im Berufsleben stehenden Zielpersonen täglich berufsbezogenes Gruppen- und Führungsverhalten zeigen, wurde dieses mit Hilfe eines 360°-Feedbacks erhoben. Obwohl das 360°-Feedback aufgrund seines multiperspektivischen Ansatzes im Allgemeinen als objektives Einschätzungsinstrument gilt (Fleenor et al., 2008; Rao & Rao, 2005), wurde zusätzlich als hartes Datum die tatsächliche berufliche Position der Zielpersonen erhoben. So konnte einerseits analysiert werden, inwiefern ein auf 20-sekündigem tonlosen Videomaterial beruhender Eindruck mit der Einschätzung durch Vorgesetzte, Kollegen und – wenn vorhanden – Mitarbeitern übereinstimmt und ob intuitive Beurteiler Mitarbeiter von Führungskräften unterscheiden können.

Vor der Analyse der Einschätzungsgüte wurde jedoch zunächst die Frage nach dem Urteilsprozess beleuchtet. Bestätigt wurden in diesem Kontext zunächst die Hypothesen der „what is beautiful is good“-Heuristik (Dion et al., 1972; Langlois et al., 2000) und des „think manager - think male“-Stereotyps (Schein, 1973, 1975; Sczesny, 2003b; Sczesny et al., 2006). Beide hatten jeweils direkte Auswirkungen auf den Wahrnehmungsprozess. Männlichen sowie attraktiveren Zielpersonen wurde eine höhere Führungsstärke zugeschrieben. Dieser Effekt zeigte sich erneut aufgrund eines Suppressor-Effektes von Attraktivität jedoch erst auf multivariater Ebene.

Die beiden anderen Hypothesen konnten dagegen nicht bestätigt werden. Weder veränderte sich die aufgeklärte Varianz des Antwortverhaltens der Beurteiler unter Einbezug der Interaktion zwischen Geschlecht und Attraktivität wie es das „Lack-of-fit“-Modell (Heilman, 1983; Heilman & Stopeck, 1985) vorausgesagt hatte, noch fanden sich kurvilineare Zusammenhänge zwischen der Attraktivität der Zielpersonen und zugeschriebener Persönlichkeitsstärke bei der gleichgeschlechtlichen Bewertung wie es der „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007) prognostiziert hatte.

Bezogen auf die Frage nach der Genauigkeit der Einschätzung von Führungsstärke zeigte sich zunächst auf bivariater Ebene ein signifikanter Zusammenhang zwischen intuitiven Beurteilern und dem 360°-Feedback. Doch auch hier zeigte sich, dass der moderate Zusammenhang von .29 auf multivariater Ebene geringer wurde und nunmehr lediglich einen marginal signifikanten Einfluss besaß. Ein Grund für diesen Effekt mag in der starken Nutzung des Geschlechts- und des Attraktivitätsstereotyps liegen. Diese beiden Merkmale zusammen erklärten etwa 53% des Antwortverhaltens der Intuitivbeurteiler, das 360°-Feedback dagegen etwa 3%. Obwohl auf der Suche nach dem „Körnchen Wahrheit“ bezogen auf das Geschlecht kein Unterschied im 360°-Feedback gefunden werden konnte, wurde attraktiveren Zielpersonen im 360°-Feedback tatsächlich eine höhere Führungsstärke zugeschrieben.

Offen bleibt hier die Frage, ob attraktivere Zielpersonen tatsächlich mehr Führungsverhalten zeigen oder ob auch die Befragten im Arbeitsumfeld der Zielpersonen stereotypen Verzerrungen unterlagen.

Bezogen auf die erste mögliche Überprüfung der Genauigkeit der Personenwahrnehmung lässt sich somit festhalten, dass die auf bivariater Ebene gefundene Übereinstimmung zwischen intuitiv zugeschriebener und tatsächlicher Führungsstärke zumindest zum Teil auf die Nutzung des Attraktivitätsstereotyps zurückzuführen war, welches eine Entsprechung im 360°-Feedback fand und somit einen validen Hinweis darstellte. Auch konnten die intuitiven Beurteiler zumindest zu einem gewissen Grad die tatsächliche Führungsstärke erkennen. Auch wenn dieser Effekt auf multivariater Ebene lediglich marginal signifikant wurde, ist es dennoch erstaunlich, dass die intuitiven Beurteiler, die die jeweilige Zielperson lediglich für 20 Sekunden ohne Ton alleine an einem Tisch sitzen und in die Kamera sprechen sahen, zu einem gewissen Grad mit dem Arbeitsumfeld dieser Zielperson übereinstimmten.

Bezogen auf das 360°-Feedback als Referenzmaß für die tatsächliche Führungskraft der Zielpersonen fällt die geringe Übereinstimmung zwischen den Einschätzungen durch Vorgesetzte, Kollegen und Mitarbeiter auf. Hier wurde lediglich der Zusammenhang zwischen den Einschätzungen der Vorgesetzten und der Kollegen einer Zielperson signifikant, alle anderen Perspektiven standen in keinem Zusammenhang. An dieser Stelle ist zu bemerken, dass die Rückmeldungen über die Zielpersonen analog zu den vorherigen Studien auf der Skala Persönlichkeitsstärke erfolgten, die angelegt ist, Führungskraft als Persönlichkeitsmerkmal zu erfassen. Auch wenn Persönlichkeit nicht mit konkretem Verhalten in einer Situation gleichzusetzen ist, sondern sich auf die Organisation und Struktur von Verhalten bezieht (Hodapp & Rohrmann, 2006), so sind es dennoch Persönlichkeitsmerkmale, die dem Verhalten Konsistenz und Stabilität verleihen.

Auf der anderen Seite steht das Argument, es sei nicht verwunderlich, dass eine Führungskraft von ihren Vorgesetzten anders bewertet wird als von ihren Mitarbeitern, da gerade diese beiden Gruppen die Person in maximal unterschiedlichen Rollen erleben. So konnte die Forschung gleichsam nachweisen, dass Interaktionseffekte zwischen Dispositionen und Situationen ebenfalls ein gängiger und alltäglicher Ausdruck von Persönlichkeit sind (Mischel, 2004). Auf der Verhaltensebene äußern sich unterschiedliche situative Anforderungen im Zusammenwirken mit den personalen Eigenschaften als ein Muster stabiler Variabilitäten, das heißt in einem individuell spezifischen Muster von Wenn (Situation)– Dann (Verhalten)– Verknüpfungen (Kammrath, Mendoza-Denton & Mischel, 2005).

Insofern kann der Befund der geringen Übereinstimmung zwischen den verschiedenen Feedbackquellen im 360°-Feedback auch als Bestärkung der Notwendigkeit eines multiperspektivischen Ansatzes gewertet werden. Durch die Mittelung der Einzeleinschätzungen sollte somit analog zu der „Weisheit der Vielen“- Logik (Surowiecki, 2005)

die Anforderungen der Situation eliminiert werden und lediglich die Information beziehungsweise das Merkmal selbst übrig bleiben.

Bei Betrachtung der Gesamtbewertungen im 360°-Feedback fiel andererseits auf, dass Führungskräften übereinstimmend mit den Ergebnissen von Noelle-Neumann (1983) eine höhere Persönlichkeitsstärke attestiert wurde als Mitarbeitern ohne Führungsverantwortung. Dieses Resultat unterstreicht die externe Validität der Skala und die Adäquatheit ihres Einsatzes im Rahmen dieser Fragestellung.

Doch selbst wenn man den geringen perspektivischen Konsens innerhalb des 360°-Feedback als Einschränkung der Ergebnisse betrachten mag, so förderte doch die zweite Analyse der Genauigkeit in dieser Studie erstaunliche Ergebnisse ans Licht.

Hier zeigte sich, dass die Beurteiler intuitiv Mitarbeiter von Führungskräften unterscheiden konnten. Darüber hinaus nutzten sie das Geschlecht und die Attraktivität der Zielpersonen als Hinweise auf die vermutete berufliche Position. Doch auch dies stellte eine sinnvolle Herangehensweise dar. Sowohl männliche als auch attraktivere Zielpersonen waren mit einer höheren Wahrscheinlichkeit in der Gruppe der Führungskräfte anzufinden, wobei der Einfluss der Attraktivität dabei lediglich marginal signifikant war.

Auffällig ist, dass die Ergebnisse auf dieser Dimension im Gegensatz zu allen bisher gefundenen Ergebnissen stehen. Nicht nur nutzten die Beurteiler Stereotype als sparsame und (teilweise) sinnvolle Heuristiken, darüber hinaus erkannten sie zudem den sozialen Status und konnten damit insgesamt die Zielpersonen besser als zufällig genau einschätzen.

7. Diskussion der empirischen Befunde

„Die Welt, mein Sohn“ erklärt im Aquarium der Vaterfisch seinem Filius, „ist ein großer Kasten voller Wasser!““

Richard David Precht (2007, S.28)

Dieser Vaterfisch, so scheint es, verfügt über eine relativ eingeschränkte Weltsicht. So glaubt er doch tatsächlich, die Welt bestünde nur aus einem Kasten voller Wasser, aus dem, was er sieht und fühlt. Dabei begreift er nicht dass, nur weil er die Welt aus seiner persönlichen Perspektive wahrnimmt, dies nicht unbedingt bedeutet, dass die Welt wirklich so ist. Er erkennt nicht die wahre Natur der Welt, sondern bleibt in seiner subjektiven Sichtweise behaftet.

Recht einfältig, diese Fische. Doch sind wir Menschen in dieser Hinsicht so anders?

Aus dem theoretischen Überblick zu Beginn dieser Arbeit wurde deutlich, dass es Wissenschaftlern nur allzu oft misslingt, über den Rand ihres persönlichen Aquariums hinwegzublicken und die Welt nicht nur durch die Brille ihres Forschungsparadigmas zu betrachten. An dieser Stelle geht die vorliegende Forschungsreihe über die im Bereich Personenwahrnehmung führenden Forschungstraditionen hinaus, da sie diese kombiniert und ihre Vorhersagen gegeneinander testet.

Es wurden vier Studien entworfen, die gleichsam sowohl den Prozess als auch das Ergebnis der Wahrnehmung von Führungsstärke beleuchteten. Mit ihrer Hilfe sollten Antworten auf die Fragen gefunden werden, ob Führungsstärke intuitiv erkennbar ist und an welchen Merkmalen sich intuitive Beurteiler bei diesem Zuschreibungsprozess orientieren.

Dabei wurde grundsätzlich angenommen, dass sowohl das Geschlecht als auch die Attraktivität der Zielpersonen einen Einfluss auf die Zuschreibung von Führungsstärke haben. Über die Richtung beziehungsweise das Zusammenspiel dieser beiden Merkmale wurden

aus den verschiedenen Forschungsrichtungen konkurrierende Hypothesen aufgestellt, welche in jeder Studie auf ihr Zutreffen hin überprüft wurden.

In einer Linie mit der sozialen Kognitionsforschung besagte die erste Hypothese, dass Beurteiler intuitiv Männern und attraktiveren Zielpersonen eine höhere Führungsstärke zuschreiben. Damit würden sie einerseits der „what is beautiful is good“-Heuristik folgen, nach welcher schönen Menschen eine Vielzahl positiver Attribute zugesprochen werden wie höherer Berufserfolg, mehr berufliche, akademische und interpersonelle Kompetenz, etc. (Dion et al., 1972; Feingold, 1992; Langlois et al., 2000). Gleichzeitig würden sie dem „think manager – think male“-Stereotyp unterliegen, welches besagt, dass erfolgreiche Führungskräfte eher mit maskulinen Eigenschaften beschrieben werden und diese auch als notwendig für eine Führungsposition erachtet werden (Eagly & Sczesny, in Druck; Powell et al., 2002; Schein, 1973, 1975; Sczesny et al., 2006).

Neben der Annahme eines direkten Einflusses beider Effekte wurde aus dem „Lack-of-fit“-Modell die Hypothese aufgestellt, dass eine Interaktion zwischen Geschlecht und Attraktivität auftritt, insofern als dass männlichen attraktiven Zielpersonen eine hohe, weiblichen attraktiven Zielpersonen dagegen eine niedrigere Führungsstärke zugeschrieben wird. Grundlage des „Lack-of-fit“-Modells ist einerseits die maskuline Typisierung von Führungspositionen und andererseits der Befund, dass Attraktivität bei Frauen mit Femininität einhergeht. Diese beiden Faktoren kombinierend geht das Modell davon aus, dass bei attraktiven Frauen im Allgemeinen die geringste Passung mit den Anforderungen einer Führungsposition wahrgenommen wird (Heilman, 1983; Heilman & Saruwatari, 1979; Heilman & Stopeck, 1985).

Die dritte Hypothese entstammte der evolutionspsychologischen Forschung zum so genannten „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007), der voraussagt, dass erstens attrakti-

ven Personen des anderen Geschlechts eine höhere Führungsstärke zugeschrieben wird (Glorifizierungskomponente), während zweitens gleichgeschlechtlichen attraktiven Kandidaten eine niedrigere Führungsstärke zugeschrieben wird als Zielpersonen mittlerer Attraktivität (Abwertungskomponente). Bei der Beurteilung gleichgeschlechtlicher Zielpersonen wurde somit ein kurvilinearere Zusammenhang zwischen der Attraktivität der Zielpersonen und der ihnen zugeschriebenen Führungsstärke erwartet. Zielpersonen von durchschnittlicher Attraktivität sollten hier bevorzugt bewertet werden.

Bezogen auf die Frage nach der Genauigkeit der Wahrnehmung von Führungsstärke konnte aus der sozialen Kognitionsforschung keine eindeutige Hypothese abgeleitet werden, da diese sich traditionellerweise stärker mit dem Prozess als mit dem Ergebnis von Personenwahrnehmung beschäftigt. Durch die Fokussierung dieses Forschungszweiges auf Abweichungen und Fehler im Denkprozess könnte allerdings der Schluss gezogen werden, dass Genauigkeit in der Personenwahrnehmung kaum zu erreichen ist, wenn Beurteiler ihre Einschätzung auf Stereotype basieren. Insofern könnte die Hypothese lauten, dass die Einschätzung der intuitiven Beurteiler verzerrt sein sollte.

Aus der Hypothese des Persönlichkeitseinschätzungs-Instinktes (Haselton & Funder, 2006) dagegen wurde die Hypothese abgeleitet, dass die intuitiven Beurteiler die Führungsstärke besser als zufällig erkennen sollten. In diesem Rahmen wurde auf die Theorie der simplen und sparsamen Heuristiken der ABC Forschergruppe um Gerd Gigerenzer verwiesen, nach der vermutet wurde, dass die intuitiven Beurteiler Heuristiken und Stereotype verwenden, deren Nutzung sie jedoch schlussendlich zu dem gewünschten Ziel führt. Dies wäre dann der Fall, wenn die Nutzung spezifischer Stereotype in diesem Falle eine sinnvolle Herangehensweise darstellen würde, beispielsweise wenn sie ein „Körnchen Wahrheit“, das heißt eine Entsprechung in der Wirklichkeit, aufweisen würden.

Jede der vier durchgeführten Studien prüfte diese Hypothesen auf ihren Gehalt und beleuchtete darüber hinaus systematisch eine andere Facette der allgemeinen Fragestellung.

Studie 1 diente dazu, explorativ zu untersuchen, ob die Wahrnehmung von Führungsstärke tatsächlich durch das Geschlecht und die Attraktivität der Zielpersonen beeinflusst wird. Darüber hinaus sollte eine erste Überprüfung der drei konkurrierenden Hypothesen über die Art und Weise des Einflusses dieser beiden Variablen stattfinden. Hierbei lag der Fokus ausschließlich auf dem Wahrnehmungsprozess, während die Genauigkeit der Wahrnehmung von Führungsstärke zunächst außer Acht gelassen wurde.

Diese Frage wurde erstmalig mit Studie 2 aufgegriffen. Zur Ermittlung der Wahrnehmungsgenauigkeit wurden die intuitiven, auf 20-sekündigen tonlosen Videosequenzen basierenden „Thin slices“-Bewertungen, mit einer atomistischen und objektiven Auszählung von führungsstarkem Verhalten der Zielpersonen verglichen. Diese Verhaltensauszählung basierte auf einer Beobachtung der Zielpersonen während einer gruppendynamisch orientierten Konstruktionsübung. Dabei wurden ihre Verhaltensweisen von geschulten Beobachtern anhand von deduktiv aus der Forschung zur charismatischen Führung (Yukl, 2002) definierten, verhaltensorientierten Dimensionen ausgezählt und dokumentiert. So konnte neben der Analyse des Zuschreibungsprozesses auch die Wahrnehmungspräzision untersucht werden.

In Studie 3 wurde untersucht, ob und wie sich Erfahrung und Expertise bei der Personenbeurteilung auf die intuitive Zuschreibung von Führungsstärke auswirken. Zu diesem Zweck wurden zwei Stichproben rekrutiert, von denen eine bisher kaum berufliche Erfahrung mit Bewertungssituationen hatte (Laien), die andere dagegen aus beruflichen Gründen regelmäßig Beurteilungen fremder Personen vornahm (Experten). Auch hier wurde die Wahrnehmungspräzision über den Vergleich der intuitiven Einschätzung von Führungsstärke mit dem

gezeigten Führungsverhalten gemessen. Als Stimulusmaterial diente eine zufällige Auswahl aus dem im Rahmen von Studie 2 generierten Stimulusmaterial.

In Studie 4 wurde eine strukturell unterschiedliche Zielpersonengruppe rekrutiert. Diese Zielpersonen waren im Vergleich zu den zuvor gefilmten Studenten älter und standen bereits im Berufsleben. Diese Tatsache erlaubte die Generierung von zweierlei unterschiedlichen Referenzkriterien für Führungsstärke. Zum einen wurde ein 360°-Feedback erhoben, welches neben einer hohen Objektivität durch die Integration mehrerer Einschätzungen aus verschiedenen Perspektiven den weiteren Vorteil hat, dass es auf dem tatsächlichen Verhalten einer Person im tagtäglichen Berufsleben basiert. Zusätzlich wurde die tatsächliche berufliche Position der Zielpersonen als kondensiertes Maß für Führungsstärke erhoben. Anschließend bewerteten Intuitivbeurteiler sowohl die Führungsstärke als auch die berufliche Position der Zielpersonen. Damit diente Studie 4 der Erweiterung des Geltungsbereiches der bisherigen Ergebnisse, indem die bislang untersuchten Hypothesen an einer Stichprobe mit unterschiedlicher Altersstruktur, beruflichem Hintergrund und einer anders gemessenen Führungsstärke, untersucht wurden.

7.1. Was bisher geschah... Zusammenfassung und Interpretation der Ergebnisse

Insgesamt dienten diese Studien der Überprüfung der aufgestellten Hypothesen zu den forschungsleitenden Fragen, erstens ob Führungsstärke intuitiv erkennbar ist und zweitens, welchen Stellenwert die Faktoren Geschlecht und Attraktivität der Zielpersonen dabei besitzen. Im Folgenden sollen die empirischen Ergebnisse der vier Studien in Bezug auf diese beiden Fragen dargestellt werden.

7.1.1.Zur Nutzung von Geschlecht und Attraktivität bei der intuitiven Zuschreibung von Führungsstärke

Dass sowohl das Geschlecht als auch die Attraktivität der Zielpersonen einen Einfluss auf die Wahrnehmung von Führungsstärke besitzen, konnte konsistent in allen vier Studien gezeigt werden.

Insgesamt betrachtet wurde attraktiveren Personen eine höhere Führungsstärke zugesprochen. Dies entspricht der aufgestellten Hypothese über die Nutzung der „what is beautiful is good“- Heuristik (Dion et al., 1972; Feingold, 1992; Langlois et al., 2000). Die Attraktivität der Zielpersonen alleine konnte dabei in den verschiedenen Studien zwischen 21% und 45% des Antwortverhaltens der Intuitivbeurteiler erklären.

Darüber hinaus zeigte sich, dass Männern entsprechend des „think manager – think male“- Phänomens (Powell et al., 2002; Schein, 1973; Sczesny et al., 2006) eine höhere Führungsstärke zugeschrieben wurde. Die einzige Ausnahme bildet hier die Studie 3, in der das Geschlecht als Orientierungspunkt im Zuschreibungsprozess bei den studentischen Beurteilern keinen Einfluss, bei den professionellen Beurteilern lediglich einen marginal signifikanten Einfluss hatte. Da jedoch bei den gleichen Zielpersonen in Studie 2 die Nutzung des Geschlechtsstereotyps nachgewiesen wurde, ist dieses Ergebnis vermutlich der geringeren Stichprobengröße zuzuschreiben (40 Zielpersonen statt 80 in Studie 2). Mit Ausnahme von Studie 3 erklärte das Geschlecht der Zielpersonen zwischen 7% und 8% der intuitiven Urteile. Insofern kann festgehalten werden, dass beide Merkmale jeweils unabhängig voneinander einen direkten Einfluss auf die Zuschreibung von Führungsstärke hatten.

Bezogen auf den derzeitigen Forschungsstand bedeuten diese Ergebnisse, dass beide Stereotype erneut als repliziert gelten können. Zu den vielen positiven Eigenschaften, die schönen Menschen zugeschrieben werden, kann nun auch Führungsstärke hinzugezählt werden. Des Weiteren scheint sich das „think manager – think male“-Phänomen insofern noch nicht gewandelt zu haben, als dass Män-

nen heute grundsätzlich mehr Führungsstärke zugetraut wird als Frauen.

Mit der Bestätigung dieser ersten, aus der kognitiven Sozialpsychologie heraus aufgestellten Hypothese geht die Widerlegung der anderen beiden konkurrierenden Hypothesen einher.

Entsprechend der aus dem „Lack-of-fit“-Modell (Heilman, 1983; Heilman & Stopeck, 1985) abgeleiteten Hypothese war vermutet worden, dass eine Interaktion zwischen dem Geschlecht und der Attraktivität der Zielpersonen einen Einfluss auf die Wahrnehmung von Führungsstärke hat. Allerdings führte der jeweils zu diesem Zweck gebildete Interaktionsterm in keiner der vier Studien zu einer Erhöhung der aufgeklärten Varianz. Somit kann festgehalten werden, dass in den vorliegenden Studien in Bezug auf die Zuschreibung von Führungsstärke Attraktivität keinen Nachteil für Frauen darstellte und die Vorhersage des „Lack-of-fit“-Modells nicht bestätigt werden konnte.

Ein zunächst vermuteter Grund für dieses Ergebnis bezog sich auf eine kulturell bedingte Veränderung des Merkmals Femininität. So gilt innerhalb des „Lack-of-fit“-Modells als Ursache für die schlechtere Bewertung der Führungsstärke attraktiver Frauen nicht deren Attraktivität per se, sondern der allgemeine Zusammenhang von Attraktivität und Femininität bei Frauen (Gillen, 1981). Kulturell bedingte Sichtweisen auf Attraktivität beziehungsweise Anziehungskraft können sich jedoch ändern, wenn es zu Veränderungen in der Gesellschaft kommt. Verschiedene Untersuchungen im deutschsprachigen und englischsprachigen Raum machen deutlich, dass maskuline Eigenschaften wie Entschlossenheit, Ehrgeiz, Intelligenz und Selbstsicherheit heute stärker zu den sozial erwünschten Eigenschaften einer Frau zählen als noch in den 80er Jahren und dass Frauen auf deskriptiver Ebene vermehrt auch mit maskulinen Charakteristiken beschrieben werden (Diekmann et al., 2005; Swazina et al., 2004). Unabhängig von den rein physischen Merkmalen von Schönheit, mag

sich die Sicht auf weibliche Attraktivität insofern verändert haben, als dass diese heute weniger eindeutig mit Femininität einhergeht.

Allerdings wurde zur Überprüfung dieses Erklärungsversuches in ergänzenden Untersuchungen für die verwendeten Zielpersonen zusätzlich deren Maskulinität beziehungsweise Femininität erhoben. Aufgrund dieser Daten wurde diese mögliche Erklärung der Ergebnisse verworfen, da die Daten starke Zusammenhänge zwischen Femininität und Attraktivität bei weiblichen Zielpersonen und zwischen Maskulinität und Attraktivität bei männlichen Zielpersonen verdeutlichten.

Eine weitere Möglichkeit, warum die Vorhersagen des „Lack-of-fit“-Modells in den vorliegenden Studien nicht bestätigt werden konnten, liegt in der unterschiedlichen Untersuchungsmethodik. Bisher in diesem Rahmen durchgeführte Studien (Heilman & Sarutawari, 1979; Heilman & Stopeck, 1985) waren Szenario-Studien, die einen konkreten, berufsbezogenen Kontext vorgaben. So könnte es sein, dass in diesen Studien die wahrgenommene Diskrepanz zwischen den Eigenschaften einer attraktiven Frau und den Anforderungen einer allgemein als maskulin charakterisierten Führungsposition durch Vorgabe einer konkreten Berufsszenerie besonders salient war. Im Unterschied zu diesen Studien war die Beurteilung in den vorliegenden Untersuchungen frei von einem berufsbezogenen Kontext, so dass die Führungsposition als Bezugspunkt nicht vorhanden war und somit ein Bewertungsmuster im Sinne des „Lack-of-fit“-Modells ausblieb.

Auch die dritte Hypothese, das Vorliegen eines „sexual attribution bias“ (Agthe et al., 2008; Försterling et al., 2007), konnte in keiner der vier Studien bestätigt werden. So sollten entsprechend dieser Vermutungen weibliche Beurteiler die höchsten Werte für Führungsstärke an attraktive männliche Zielpersonen und an weibliche Zielpersonen durchschnittlicher Attraktivität vergeben. Analog dazu sollten männliche Beurteiler attraktive weibliche Zielpersonen und

mittelattraktive männliche Zielpersonen hoch bewerten. Mit anderen Worten sollte es bei gegengeschlechtlicher Einschätzung zu einem linearen Zusammenhang zwischen Attraktivität und Führungsstärke kommen, da eine glorifizierende Sichtweise potenzieller Partner eine wichtige evolutionäre Funktion in Bezug auf Partnersuche und Bindung in einer Partnerschaft erfüllt. Bei gleichgeschlechtlicher Einschätzung sollte sich ein kurvilinearere Zusammenhang zwischen Attraktivität und Führungsstärke aufzeigen lassen, da es im intrasexuellen Wettbewerb um begehrten Partner für das eigene Selbstvertrauen sinnvoll ist, Konkurrenten subjektiv abzuwerten.

Der vorhergesagte lineare Zusammenhang konnte gefunden werden, genau so wie es auch die „what is beautiful is good“-Heuristik vorausgesagt hätte. Allerdings erwies sich die Annahme eines kurvilinearen Zusammenhanges in den gleichgeschlechtlichen Bewertungskombinationen in keiner der Studien als angemessener als die Annahme eines linearen Zusammenhangs.

Auch das Ausbleiben dieses Attributionsmusters liegt vermutlich im Unterschied in Bezug auf den vorgegebenen Kontext beziehungsweise in der Einschätzungsaufgabe der Beurteiler im Rahmen dieser Studien im Vergleich zu den vorher stattgefundenen Untersuchungen. Zum einen wurde in den vorliegenden Studien kein berufsbezogener Kontext angegeben. Vor allem jedoch unterschied sich die Aufgabe der Beurteiler hier deutlich von den anderen Studien.

In der vorliegenden Studienreihe war die Aufgabe jeweils die Einschätzung eines Persönlichkeitsmerkmals. Vorhergehende Studien (Agthe et al., 2008; Försterling et al., 2007; Luxen & Van De Vijver, 2006) arbeiteten mit verschiedenen anderen Aufgaben, wie Erfolgs- und Misserfolgsattributionen, auf die Zielperson gerichtete Emotionen, mit Entscheidungen darüber, welcher Kandidat eine bestimmte Stelle oder ein Stipendium erhalten soll oder mit Angaben darüber, mit welcher der Zielpersonen man gerne befreundet wäre oder zusammenarbeiten würde. Mit anderen Worten lässt sich der „sexual

attribution bias“ dann finden, wenn man Menschen fragt, wen sie gerne im beruflichen oder im privaten Bereich um sich haben würden oder wie sie (Miss-) Erfolg bei ihren Mitmenschen erklären.

Evolutionär gesehen kann sich eine solche Verzerrungstendenz wie der „sexual attribution bias“ nur halten, wenn die Kosten der Fehlwahrnehmung (Glorifizierung beziehungsweise Abwertung) geringer sind als der Nutzen der Genauigkeit.

In der Forschung zum „sexual attribution bias“ zeigt sich bislang, dass Menschen lieber mit attraktiven gegengeschlechtlichen und mittelattraktiven gleichgeschlechtlichen Personen zusammen sind, lieber mit ihnen zusammen arbeiten und lieber mit ihnen befreundet wären (Agthe et al, 2008; Luxen & Van De Vijver, 2006). Evolutionär gesehen macht dies durchaus Sinn. Durchschnittlich attraktive Konkurrenten verschlechtern die eigene Position im intrasexuellen Wettbewerb nicht unbedingt, dagegen erhöht der Umgang mit vielen attraktiven Partnern die Chance, einen passenden Partner zu finden.

Auch ein „sexual attribution bias“ in Bezug auf die Attribution von Erfolg oder Misserfolg mag evolutionär gesehen sinnvoll sein. Beispielsweise könnte es das eigene Selbstvertrauen und die Wettbewerbsmotivation verringern, würde man den Erfolg von Konkurrenten mit deren hohen Fähigkeiten erklären statt mit Zufall. Hier könnte eine genaue Einschätzung dazu führen, sich selbst im Wettbewerb zurückzunehmen und damit würde die Wahrscheinlichkeit der eigenen Reproduktion sinken. Andersherum kann es sinnvoll sein, den Erfolg attraktiver potenzieller Partner auf Fähigkeiten zu attribuieren als auf Zufall oder Glück, da sie auf diese Weise noch begehrenswerter erscheinen (Försterling et al., 2007). Dadurch erhöht sich erstens die Chance, eine Person anziehend zu finden und zweitens ist die Wahrscheinlichkeit höher, dass Menschen sich um die Gunst der glorifizierten potenziellen Partner bemühen und mit diesen anschließend eine dauerhafte Beziehung beginnen und aufrechterhalten.

Diese Logik ist jedoch nicht unbedingt auf die Wahrnehmung von Führungskraft übertragbar. Bedenkt man die zu Beginn verdeutlichte Relevanz von sozialem Status und dessen Zusammenhänge mit Führungskraft, so wird deutlich, dass die Kosten einer Fehlwahrnehmung des sozialen Status anderer deutlich höher waren und der Selektionsdruck somit vermutlich eher auf Genauigkeit lag. In diesem Kontext legen verschiedene Forschungsergebnisse nahe, dass man sowohl den relativen sozialen Status anderer Personen als auch seinen eigenen adäquat einschätzen kann (Anderson et al., 2006; Kalma, 1991). Auch in den vorliegenden Studien zeigte sich, dass der soziale Status der Zielpersonen zutreffend erkannt wurde.

Zusammenfassend lässt sich festhalten, dass von den drei konkurrierenden Hypothesen die „what is beautiful ist good“- und die „think manager – think male“- Heuristik bestätigt werden konnten, während die Vorhersagen des „Lack-of-fit“- Modells und des „sexual attribution bias“ keine Bestätigung fanden.

7.1.2.Zur Genauigkeit der Wahrnehmung von Führungskraft

Bezogen auf die Frage, ob Menschen die Führungskraft anderer intuitiv erkennen können, zeigten sich über die verschiedenen Studien hinweg gemischte Ergebnisse.

In allen Studien zeigte sich auf bivariater Ebene ein Zusammenhang zwischen den intuitiven Bewertungen auf Basis der jeweils 20-sekündigen tonlosen Videosequenzen und dem jeweiligen Vergleichskriterium für Führungskraft. Das Referenzkriterium zur Feststellung der Genauigkeit der Personenwahrnehmung bestand in Studie 2 und 3 in tatsächlich gezeigtem führungsstarken Verhalten, in Studie 4 im 360°-Feedback.

Die auf bivariater Ebene gefundenen Zusammenhänge bewegten sich zwischen .23 in Studie 2 bis .29 in Studie 4. Hypothesenkonform waren die Einschätzungen somit besser als bei rein zufälligen Urteilen zu erwarten. Bedenkt man, dass die intuitiven Beurteiler die

Zielpersonen in einer Situation ohne sozialen Kontext, ohne das verbale Verhalten und nur für einen sehr kurzen Zeitraum sahen, ist es erstaunlich, dass sie überhaupt etwas erkennen konnten.

Weiterhin zeigte sich, dass die durchschnittlich gefundenen Zusammenhänge von .27 zwischen den Urteilen der Laien und den entsprechenden Referenzkriterien dem von Ambady und Kollegen (2000) durchschnittlich gefundenen Zusammenhang von $r = .27$ entspricht und somit als ein repräsentatives Ergebnis für ein „Thin slices“-Experiment anzusehen ist. Würden keine weiteren Analysen folgen – der Regelfall im „Thin slices“-Paradigma – so wäre die Aussage der vorliegenden Studien gewesen, dass erste Eindrücke über die Führungsstärke fremder Personen zu einem gewissen Grad von Genauigkeit zeugen.

Als ein Kritikpunkt an der „Thin slices“-Forschung wurde allerdings zu Beginn geäußert, dass diese über die Feststellung eines bivariaten Zusammenhangs selten hinausgeht. Analysiert man die bivariaten Zusammenhänge auf Individualebene, so zeigt sich, dass ein Teil dieses Genauigkeitseffektes auf die Aggregation der Einzelwerte zurückzuführen ist. Auf Einzelebene berechnet ist die Beziehung zwischen den jeweiligen Referenzwerten und den intuitiven Urteilen weitaus geringer, im Durchschnitt lediglich .14.

Auch in der multivariaten Datenanalyse zeigte sich ein geringer Genauigkeitseffekt. Der Einfluss des Referenzkriteriums auf die intuitive Wahrnehmung von Führungsstärke wurde nur in einem Fall in Studie 4 marginal signifikant, in allen anderen Fällen hatte es keinen Einfluss. In Studie 2 und 3 zeigte sich stattdessen, dass die intuitiven Beurteiler sich an den Merkmalen Attraktivität und Geschlecht (mit Ausnahme der studentischen Beurteiler in Studie 3) orientierten, darüber hinaus jedoch die wahre Führungsstärke der Zielpersonen nicht erkannten. Durch den Einfluss von Geschlecht und Attraktivität auf das Intuitivurteil verringerten sich die auf bivariater Ebene gefundenen signifikanten Zusammenhänge auf multivariater Ebene

auf durchschnittlich .13 und konnten somit lediglich 2% des Antwortverhaltens der intuitiven Beurteiler erklären.

Schließlich wurde in Studie 2 und 3 das Referenzkriterium auf das Vorliegen eines „Körnchen Wahrheits“ überprüft. Geschlecht und Attraktivität hatten jedoch keinen Einfluss auf die Skala Führungsverhalten, weder erreichten Männer höhere Werte auf der Skala Führungsverhalten noch attraktivere Zielpersonen. Somit kann festgehalten werden, dass die Beurteiler sich zwar an Stereotypen orientierten, diese allerdings kein „Körnchen Wahrheit“ besaßen und die wahre Führungsstärke nicht erkannt werden konnte.

Als weiteres wichtiges Ergebnis aus Studie 3 ist zu betonen, dass die Experten in ihrer Einschätzungsgüte die Laien nicht übertrafen. Dieses Ergebnis widerspricht der Lernperspektive der Intuition und deutet an, dass Personenbeurteilungen durch mehr Erfahrung und Übung nicht unbedingt besser werden.

Fasst man die Ergebnisse bis zu diesem Punkt zusammen, so müsste nahezu uneingeschränkt der sozialen Kognitionsforschung zugestimmt werden. Aus dieser Perspektive heraus sind die Ergebnisse auch gut zu erklären: Führungsstärke ist ein komplexes soziales Charakteristikum, auf das auf der Basis einer ausdrucks- und kontextarmen Situation geschlossen werden soll. Aus Mangel an Informationen sind die Beurteiler gezwungen, sich an leicht wahrnehmbaren und schnell verwertbaren Hinweisen zu orientieren. Als Basiskategorie bietet sich das Geschlecht in diesem Rahmen geradezu an, ebenso wie die „what is beautiful is good“-Heuristik als geläufiges und reichhaltiges Stereotyp. So kommt es, dass die Beurteiler sich ausschließlich auf solche Merkmale konzentrieren, ohne dabei die tatsächlichen Eigenschaften des Gegenübers zu erkennen. Da beide Stereotype in diesem Fall kein „Körnchen Wahrheit“ besitzen, fällen die Beobachter ein Urteil, das nicht der Realität entspricht und somit als verzerrt anzusehen ist. Die Ergebnisse von Studie 4 ändern diese Aussagen jedoch in einigen Punkten.

Bezogen auf das Referenzkriterium 360°-Feedback in Studie 4 zeigte sich auf der einen Seite, dass insgesamt den Führungskräften unter den Zielpersonen eine höhere Persönlichkeitsstärke zugeschrieben wurde. Dieser Befund unterstreicht die externe Validität und damit die Adäquatheit des Einsatzes dieser Skala. Auf der anderen Seite zeigte sich, dass Vorgesetzte, Kollegen und Mitarbeiter in ihrem Urteil über die Zielperson nur wenig übereinstimmten. Diese Tatsache betont einerseits, im Hinblick auf den Anspruch der Objektivität dieses Maßes, die Notwendigkeit der Aggregation dieser unterschiedlichen Perspektiven, stellt gleichzeitig jedoch die Eigenschaft von Persönlichkeitsstärke als dispositionelles Merkmal in Frage.

Bezogen auf die Beurteilungsgenauigkeit enthüllte die bivariate Analyse der Daten auch hier eine signifikante Übereinstimmung zwischen intuitiv wahrgenommener Führungsstärke und 360°-Feedback.

Das 360°-Feedback erwies sich in dieser Studie im Gegensatz zu den vorherigen Studien auch auf multivariater Ebene neben Geschlecht und Attraktivität der Zielpersonen als dritter, wenn auch lediglich marginal signifikanter Einflussfaktor auf die intuitiv wahrgenommene Führungsstärke. Insgesamt konnte das 360°-Feedback etwa 3%, das Geschlecht etwa 7% und die Attraktivität der Zielpersonen etwa 22% des Antwortverhaltens der Intuitivbeurteiler erklären.

Auch trug in dieser Studie die Nutzung des Attraktivitätsstereotyps zur Wahrnehmungsgenauigkeit bei, da attraktivere Personen auch im 360°-Feedback höhere Werte erlangten und somit dieses Stereotyp ein „Körnchen Wahrheit“ besaß. Offen bleibt hier die Frage, welche Mechanismen diesen Effekt verursachen. Einerseits ist es durchaus möglich, dass attraktivere Personen tatsächlich führungsstärker sind. Genauso sind jedoch alternative Erklärungen denkbar. So könnte sich das Arbeitsumfeld attraktiveren Personen gegenüber genau so verhalten, dass sie Führungsverhalten im Sinne einer Selbst-erfüllenden-Prophezeiung evozieren. Darüber hinaus wäre es

möglich, dass gleiches Verhalten bei attraktiveren Personen schlichtweg wohlwollender interpretiert wird.

Diese Frage ist jedoch mit den vorliegenden Studien nicht zu beantworten. Zukünftige Forschung sollte dieser Frage nachgehen und ein erhobenes 360°-Feedback über die Zielpersonen an zusätzlichen Messungen validieren. Denkbar wäre beispielsweise weitere Kennzeichen für den Erfolg einer Führungskraft heranzuziehen wie Stückzahlen oder den Erfolg der Gruppe. Weiterhin könnten die Zielpersonen zusätzlich an einer oder mehreren Assessment Center Übungen teilnehmen. Während dieser Übungen sollte das führungsrelevante Verhalten gezählt werden. Durch Vergleich der Werte im 360°-Feedback mit dem tatsächlich gezeigten Führungsverhalten kann überprüft werden, ob im 360°-Feedback besser bewertete Personen mehr Führungsverhalten zeigen. Insbesondere könnte so analysiert werden, ob attraktive Personen mehr Führungsstärke zeigen. Somit könnte gleichsam untersucht werden, ob das gefundene „Körnchen Wahrheit“ des Attraktivitätsstereotyps im 360°-Feedback tatsächliche Führungsstärke repräsentiert oder von einer kollektiven Wahrnehmungsverzerrung des Arbeitsumfeldes zeugt.

Insbesondere mit der Analyse des letzten Indikators für Wahrnehmungsgenauigkeit, dem Vergleich der tatsächlichen beruflichen Position der Zielpersonen mit der intuitiv vermuteten beruflichen Position, veränderte sich das Gesamtbild noch einmal. Hier zeigte sich, dass die Beurteiler intuitiv Führungskräfte von Mitarbeitern unterscheiden konnten. Eine zu Illustrationszwecken durchgeführte Dichotomisierung der aggregierten Antworten konnte zeigen, dass von den intuitiven Beurteilern im Durchschnitt 100% der Führungskräfte korrekt als solche identifiziert wurden sowie fast 81% der Mitarbeiter. Auf multivariater Ebene zeigte sich, dass die tatsächliche berufliche Position auch über die beiden Stereotype hinaus einen signifikanten Einfluss auf die Wahrnehmung der Intuitivbeurteiler hatte.

Offensichtlich gelang es den intuitiven Beurteilern somit, den sozialen Status der Zielpersonen einzuschätzen. Dies war aus der evolutionspsychologischen Perspektive aus der Wichtigkeit von Statusverhandlungen und dem damit verbundenen Zugang zu Ressourcen heraus vermutet worden und entspricht anderen Forschungsergebnissen, die genaue Einschätzungen des eigenen Status und des Status anderer Menschen demonstrierten (Anderson et al., 2006; Kalma, 1991).

Darüber hinaus zeigte sich, dass die Anwendung beider Stereotype im Sinne von simplen und sparsamen Heuristiken zu guten Ergebnissen führte. Dies war vor allem auf das „Körnchen Wahrheit“ zurückzuführen, da unter den Führungskräften tatsächlich mit einer höheren Wahrscheinlichkeit Männer anzutreffen waren sowie attraktivere Personen, wobei der Einfluss der Attraktivität auf die berufliche Position lediglich marginal signifikant war.

Fasst man die Ergebnisse der vorliegenden Studienreihe in Bezug auf die Genauigkeit der Wahrnehmung von Führungskraft zusammen, so könnte man argumentieren, dass die meisten Ergebnisse zeigen, dass Menschen diese Fähigkeit nicht besitzen. Allerdings stehen dieser globalen Interpretation die Ergebnisse der letzten Analyse entgegen, die zeigen konnte, dass Menschen mit ihren Einschätzungen doch nicht nur irren. Doch wie ist dies möglich angesichts der konträren Ergebnisse der vorhergehenden Studien? Auf der Suche nach möglichen Erklärungen für diese Diskrepanz ist ein möglicher Grund in den Unterschieden des letzten Indikators für Genauigkeit im Vergleich zu den zuvor verwendeten zu finden.

7.1.3.Kritische Betrachtung der verwendeten Referenzkriterien

Zum einen stellt die berufliche Position der Zielpersonen ein qualitativ anderes Kriterium dar als die Verhaltensbeobachtung und das 360°-Feedback. In der beruflichen Position kondensiert sich das gesamte berufliche Verhalten, während die zuvor genutzten Referenz-

kriterien lediglich Annäherungen an Objektivität und tatsächliche Führungskraft darstellen.

Die Beobachtung führungsrelevanten Verhaltens innerhalb einer Assessment Center Übung, wie in Studie 2 und 3 genutzt, stellt zwar eine sinnvolle, jedoch zwangsläufig reduktionistische Herangehensweise dar. Eine solche Übung ist und bleibt ein Verhaltensausschnitt von lediglich 20 Minuten. Prinzipiell ist das gezeigte Verhalten in dieser Situation zwar einer Disposition für Führungsverhalten zuzuschreiben, resultiert darüber hinaus jedoch aus vielen spezifischen situativen Faktoren wie beispielsweise die spezifische Gruppenzusammensetzung, Stimmung, Motivation, Ernsthaftigkeit, mit der die Aufgabe verfolgt wird, usw.

Darüber hinaus zeigt die Forschung, dass Männer und Frauen je nach spezieller Aufgabe mehr oder weniger Führungsverhalten zeigen. Bei aufgabenorientierten Gruppenübungen nehmen Männer im Allgemeinen eher die Führungsrolle ein, während bei gruppenorientierten Aufgaben Frauen eher die Führungsrolle einnehmen (Eagly & Karau, 2002). Durch die Kombination einer gruppenorientierten Konstruktionsaufgabe mit einem Wettbewerb (Konkurrenz mit anderen Gruppen um den höchsten, stabilsten und kreativsten Turm) wurde versucht, eine Aufgabe mit gemischtem Charakter zu kreieren, so dass die Wahrscheinlichkeit der Übernahme einer Führungsrolle für Männer wie für Frauen etwa gleich hoch war. In einer maskulinen Arbeitswelt wäre es jedoch zumindest prinzipiell möglich, dass es ein „Körnchen Wahrheit“ im Geschlechterstereotyp gibt, welches durch die spezielle Konzeption der Aufgabe in den vorliegenden Studien verborgen blieb.

Aus dieser methodischen Kritik heraus wurde in Studie 4 als Referenzkriterium ein 360°-Feedback verwandt. Grundannahme hier war erstens, dass die Einschätzung durch Vorgesetzte, Kollegen und Mitarbeiter auf Basis des tagtäglichen und tatsächlichen Verhaltens der Zielpersonen getroffen wird und daher valider sein sollte als eine

Verhaltensstichprobe innerhalb einer Assessment Center Übung. Zweitens sollte durch den multiperspektivischen Ansatz eine höhere Objektivität erreicht werden. Auf der anderen Seite sind Einflussfaktoren wie Globalurteile aufgrund von Sympathie oder Antipathie, Misstrauen, usw. nicht auszuschließen, obwohl absolute Anonymität gewährleistet wurde.

Im Gegensatz zu den beiden vorher verwendeten Referenzkriterien ist anzunehmen, dass die Mehrheit der Führungskräfte sich in ihrem gesamten beruflichen Habitus von Mitarbeitern ohne Führungsverantwortung unterscheiden und somit aus gutem Grund eine Führungsposition bekleiden. Hier bleibt somit festzuhalten, dass dem Erkennen der beruflichen Position als verdichtetes Maß für die Führungskraft einer Person ein höherer Wert zukommt als der Vorhersage von Verhalten in einer Assessment Center Gruppenübung beziehungsweise der Übereinstimmung mit einem 360°-Feedback.

Betrachtet man in diesem Rahmen die berufliche Position genauer, so fällt auf, dass sich offensichtlich hier beide Stereotype bewahrheiten. Auch in dieser Hinsicht besteht ein Unterschied zu den anderen beiden Referenzkriterien. Bei der Verhaltensbeobachtung zeigte sich keinerlei Unterschied im Verhalten in Bezug auf Geschlecht und Attraktivität der Zielpersonen. Das 360°-Feedback wurde von der Attraktivität der zu beurteilenden Personen, nicht jedoch vom Geschlecht beeinflusst. Im Gegensatz dazu hatte das Geschlecht jedoch sehr wohl einen Einfluss auf die Wahrscheinlichkeit, eine Führungskraft zu sein, ebenso wie Attraktivität die Wahrscheinlichkeit marginal signifikant erhöhte. Das in der beruflichen Position gefundene „Körnchen Wahrheit“ impliziert, dass die Nutzung beider Merkmale als Orientierung für den Schluss auf die vermutete berufliche Position zu einer guten Einschätzung führte. Doch woher kommt das „Körnchen Wahrheit“, wenn genau dies in den anderen Referenzkriterien nicht gefunden werden konnte?

Einerseits wäre es möglich, dass attraktivere Menschen und Männer im Berufsalltag und über die Zeit hinweg tatsächlich mehr, überzeugenderes oder effektiveres Führungsverhalten zeigen und aus diesem Grund eine höhere berufliche Position erreichen.

Andererseits konnten alle hier vorgestellten Studien zeigen, dass Menschen bei attraktiveren Personen und bei Männern eine höhere Führungsstärke wahrnehmen. Eine andere mögliche Erklärung für diesen Befund könnte daher aus einer Interaktion zwischen den in der Gesellschaft vorhandenen Geschlechts- und Attraktivitätsstereotypen mit Sich-selbst-erfüllenden-Prophezeiungen stammen. Der Kernpunkt der charismatischen Führungstheorie besteht darin, dass sich Menschen einer überzeugenden Führungsperson bereitwillig anschließen, sich mit ihr identifizieren und ihre Energien und Ressourcen auf die Ziele dieser Führungsperson ausrichten, wenn sie deren Fähigkeiten und Plänen respektieren (Weber, 1947; Weiss, 2003; Yukl, 2002). Sorgen nun kulturell tradierte Annahmen in Bezug auf Männer sowie insgesamt in Bezug auf attraktive Personen dafür, dass das soziale Umfeld bei diesen Personen mehr Führungsstärke beziehungsweise Charisma wahrnimmt, so werden Menschen attraktiveren Personen und Männern bereitwilliger folgen. So könnte die berufliche Position ebenso das Resultat einer Sich-selbst-erfüllenden-Prophezeiung sein.

Doch auch wenn man die Möglichkeit einer Sich-selbst-erfüllenden-Prophezeiung annimmt, bleibt eine wichtige Frage unbeantwortet. Warum wird Männern und attraktiveren Menschen mehr Führungsverhalten, ein höherer Wert im 360°-Feedback und schließlich eher eine berufliche Position mit Führungsverantwortung zugetraut?

Annahmen zur Erklärung dieses Phänomens können dem evolutionspsychologischen Ansatz entnommen werden. So könnten sowohl die Attraktivität als auch das Geschlecht früher Hinweise auf gute und effektive Führungsqualitäten dargestellt haben, heute jedoch anachronistisch sein. Da Attraktivität ein Indikator für Gesundheit

und gute Gene ist (Berry, 2000; Sugiyama, 2005), könnten zu der Zeit, in der unsere Vorfahren lebten und sich unsere heutigen psychologischen Adaptationen entwickelten, attraktivere Führende bevorzugt worden sein. Weiterhin könnten maskuline Führer wegen ihres athletischen Körperbaus und Durchsetzungsvermögens bei der Jagd und in der intrasexuellen Konkurrenz um hohen sozialen Status von der sozialen Gemeinschaft favorisiert worden sein.

Doch auch wenn man diese Argumentation gelten lässt, stellt sich die Frage, warum sich solche Stereotype bis in die heutige Zeit halten. Möglicherweise entwickelten sich auf Basis von einst guten Hinweisen auf Führungsstärke psychologische Adaptationen und gleichzeitig kulturelle Traditionen. Diese verändern sich erst dann, wenn sich Kulturen und Lebensumstände verändern. So mag das Streben nach Gleichberechtigung der Geschlechter historisch betrachtet viel zu jung sein, um substantielle Veränderungen von Stereotypen, kulturellen Einstellungen und insbesondere von zugrunde liegenden psychologischen Adaptationen zu erwirken (Richerson & Boyd, 2001; 2005). Auch heute sind viele Gesellschaften maskulin und vornehmlich von Männern dominiert. Und auch in den westlichen Industrienationen hat sich erst in den letzten Jahrzehnten die Rolle der Frau stark verändert (Diekmann et al., 2005). Da bis vor wenigen Jahren in der Berufswelt vornehmlich Männer anzutreffen waren, scheint es kaum verwunderlich, dass dieser Bereich bis heute von maskulinem Denken geprägt ist.

In einer männlich dominierten Welt waren naturgemäß auch die Führungsrollen maskulin ausgelegt, was wiederum dazu geführt haben mag, dass Männer tatsächlich die effektiveren Führungskräfte darstellten. In Situationen oder Institutionen, in denen ein autokratischer Führungsstil wie beispielsweise beim Militär erwartet wird, sind männliche Führende in der Tat erfolgreicher (Eagly, Karau & Makhijani, 1995). Der Wunsch nach Einfühlungsvermögen, Kommunikationsfähigkeiten, sozialen Kompetenzen, einem kooperativen oder

demokratischen Führungsstil bei Führungskräften erwacht aufgrund zunehmender Demokratisierung und abnehmender Hierarchien langsam aber stetig. Damit steigt auch der Erfolg von Frauen in Führungsrollen, da sie im Gegensatz zu Männern in Situationen, die einen demokratischen Führer erfordern, erfolgreicher sind (Eagly et al., 1995).

Aus diesen Überlegungen folgt auch, dass Männer und Frauen unterschiedliche Führungsqualitäten mitbringen und somit das Geschlecht weiterhin ein guter Hinweis bleibt, um Führungsstärke zu erkennen, vorausgesetzt es gelingt eine flexiblere Anwendung des Geschlechterstereotyps je nach Führungsrolle und gewünschtem Führungsstil.

Zusammenfassend könnte ein Zusammenspiel aus unserem evolutionären Erbe und überlieferten Normen der Grund sein, warum Männern eine höhere Führungsstärke zugesprochen wird, auch wenn dies im Zeitalter des Gleichberechtigungsstrebens und der weiblichen Emanzipation nicht mehr zeitgemäß erscheint. In Bezug auf die Verwendung von Attraktivität als Hinweis auf Führungsstärke kann der gleichen interaktiven Erklärungslogik gefolgt werden. So zeigen Studien wieder und wieder den massiven Halo-Effekt, den Attraktivität auslöst (Langlois et al., 2000). Aus evolutionärer Perspektive stammt dieser Effekt vornehmlich aus der Verlinkung von Attraktivität mit guten Genen beziehungsweise physischer Gesundheit und Kraft. Offensichtlich können sich Menschen diesem Ausstrahlungseffekt auch bei der Wahrnehmung von Führungsstärke nicht entziehen.

7.2. Fortsetzung folgt... Aufgaben zukünftiger Forschung

Obwohl insgesamt viele Fragen bereits beantwortet wurden, bleiben einige Rätsel ungelöst. Darunter befindet sich beispielsweise die Ungewissheit darüber, was genau die Beurteiler in diesen Studien

eigentlich erkannten. Offensichtlich konnten sie den sozialen Status der Zielpersonen zutreffend erkennen. Dies stellt ein möglicher Ansatzpunkt für zukünftige Forschung dar.

Wenn es Menschen leichter fällt, ein Ergebnis anstelle einer Persönlichkeitsvariablen vorherzusagen, könnten zukünftige Studien gezielt diesen Unterschied untersuchen. Möglich wäre, Zielpersonen erneut in einer gruppenorientierten Assessment Center Übung zu beobachten und anschließend die Teilnehmer in eine Rangfolge je nach relativem Status oder nach Beitrag zur Fertigstellung der Gruppenaufgabe zu bringen. Aufgabe der intuitiven Beurteiler wäre es dann, aus den Videosequenzen der Gruppenmitglieder diejenige Person anzuzeigen, die sich ihrer Meinung nach als Anführer der Gruppe herauskristallisiert hat oder den größten Beitrag geleistet hat.

Weiterhin ist bisher nicht deutlich geworden, wie genau die Beurteiler den sozialen Status erkennen konnten. Die Nutzung von Geschlecht und Attraktivität, inklusive des „Körnchen Wahrheits“ in diesen, reicht alleine nicht aus, um die Treffsicherheit zu erklären.

Eine Vermutung in diesem Kontext ist, dass die Beurteiler eine Art globale Selbstsicherheit oder Ausgeglichenheit wahrnahmen, die Führungskräfte möglicherweise in einem höheren Maße auszeichnet als Mitarbeiter. Um dies zu untersuchen, könnten bei den Zielpersonen entsprechende Merkmale erhoben werden, beispielsweise Selbstwirksamkeitserwartungen, Selbstwertgefühl, Selbstsicherheit, Emotionale Stabilität, usw. Eine weitere Möglichkeit besteht darin, direkt nach der Videoaufnahme die Emotionen der Zielpersonen zu messen. Damit könnte analysiert werden, wie wohl diese sich in der Aufnahmesituation fühlten, wobei vermutet würde, dass Personen, die insgesamt positivere Werte angeben, intuitiv eine höhere Führungsstärke zugeschrieben würde. Ein solches Ergebnis würde die Vermutung untermauern, dass Beurteiler sich an der ausgestrahlten Ausgeglichenheit beziehungsweise Selbstsicherheit orientieren.

Darüber hinaus könnte es möglich sein, dass die Beurteiler zwar die Führungskraft selber nicht akkurat erkennen können, diese jedoch durch die akkurate Wahrnehmung anderer Merkmale erschließen konnten. So zeigen die Resultate der eigenschaftsorientierten Ansätze der Führungsforschung verschiedene Merkmale auf, die Führungskräfte im Allgemeinen im Vergleich zu Mitarbeitern ohne Führungsverantwortung auszeichnen. Dazu gehören beispielsweise Gewissenhaftigkeit, Extraversion (Barrick, Mount & Judge, 2001), ein überdurchschnittlich hohes allgemeines Intelligenzniveau (Schmidt & Hunter, 2004), Kommunikationsgeschick, gutes Sozialverhalten und eine ausgeprägte Leistungsmotivation (Schuler, 2006). Bei einigen dieser Eigenschaften konnte bereits gezeigt werden, dass diese intuitiv erkennbar sind, so beispielsweise bei Extraversion und Intelligenz (Borkenau & Liebler, 1993; Borkenau, Mauer, Riemann, Spinath & Angleitner, 2004; Reynolds & Gifford, 2001). So könnte es sein, dass Beurteiler eher diese Merkmale zutreffend einschätzen und aus diesen heraus mit Hilfe impliziter Führungstheorien (Offermann, Kennedy & Wirtz, 1994) auf die berufliche Position der Zielpersonen schließen.

Ein dritter möglicher Ansatzpunkt zukünftiger Forschung besteht in der komplexen Natur des Konstrukts Führungskraft selber. Ein möglicher Grund, warum diese im Vergleich zu sozialem Status nicht erkennbar war, könnte darin liegen, dass sich dieses Merkmal unter Umständen erst in sozialer Interaktion manifestiert. So wäre eine Studie denkbar, in der sich die Situationen, in denen die Zielpersonen gefilmt werden, systematisch unterscheiden. So könnten in einer Bedingung die Zielpersonen alleine mit der Kamera in einem Raum sein, in der zweiten Bedingung könnte ein Versuchsleiter anwesend sein und in einer dritten Bedingung könnten die Zielpersonen in einer Interaktion zu sehen sein.

7.3. Fazit

Die vorliegende Studienreihe wurde konzipiert, um zwei zentrale Fragen zu untersuchen. Die erste Frage lautete, ob Menschen die Führungskraft anderer intuitiv erkennen können. Die zweite Frage war, ob und auf welche Art und Weise das Geschlechts- sowie das Attraktivitätsstereotyp die Wahrnehmung beeinflussen.

Die Resultate dieser Studien liefern viele Antworten. Sie zeigen, dass beide Stereotype die Wahrnehmung von Führungskraft beeinflussen. Geschlecht und Attraktivität sind zwei der Merkmale, an denen sich Menschen bei der Zuschreibung von Führungskraft orientieren. Sie schreiben Männern sowie attraktiveren Personen eine höhere Führungskraft zu.

Die Studien zeigen auch, dass Menschen die Führungskraft anderer intuitiv nicht gut einschätzen können, während sie den sozialen Status anderer offenbar erkennen.

Diese empirische Analyse der Genauigkeit von Personenwahrnehmung ermöglichte viele spannende Erkenntnisse und wirft gleichzeitig neue Fragen auf.

Woran erkennen Menschen den sozialen Status anderer? Sind attraktive Menschen tatsächlich führungsstärker? Können Menschen Führungskraft in sozialen Interaktionen erkennen? Könnte man Menschen beibringen, gute soziale Urteile zu fällen?

So bringen die vorgestellten Ergebnisse gleichsam eine weitere Erkenntnis mit sich: Es gibt noch viel zu tun.

8. Literaturverzeichnis

- Adams, H.F. (1927). The good judge of personality. *Journal of Abnormal and Social Psychology*, 22, 172-181.
- Agthe, M., Spörrle, M., Dörfler, R. & Försterling, F. (2008). Job applications, voting behavior, and university recruitment: Evidence for a multidimensional attractiveness-gender bias in selection contexts. Talk. 25th International Congress of Psychology, July 20th – 25th, 2008, Berlin (Germany).
- Allport, G.W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Ambady, N., Bernieri, F.J. & Richeson, J.A. (2000). Toward a histology of social behavior: judgmental accuracy from thin slices of the behavioral stream. In P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201-271). San Diego, CA: Academic Press.
- Ambady, N., Hallahan, M. & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, 77, 538-547.
- Ambady, N. & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 64, 431-441.
- Anderson, C., John, O. P., Keltner, D. & Krings, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology*, 81(1), 116-132.
- Anderson, C., Srivastava, S., Beer, J.S., Spataro, A.E. & Chatman, J.A. (2006). Knowing your place: Self perceptions of status in face-to-face groups. *Journal of Personality and Social Psychology*, 91(6), 1094-1110.
- Babad, E., Bernieri, F. & Rosenthal, R. (1989). Nonverbal communication and leakage in the behavior of biased and unbiased teachers. *Journal of Personality and Social Psychology*, 56(1), 89-94.
- Balyan, A.K., Kanwar, Y.P.S. & Pant, H. (2005). Identification and development of potential executives for top level positions through 360 degree feedback. In T.V. Rao & N. Chawla (pp.3-13), *360 Degree Feedback and Assessment & Development Centers*. New Delhi: Excel Books.

- Bar, M., Neta, M. & Linz, H. (2006). Very first impressions. *Emotion*, 6(2), 269-278.
- Barlow, N. (2005). *The autobiography of Charles Darwin 1809-1882*. New York, NY: W.W. Norton & Company, Inc.
- Barrick, M.R., Mount, M.K. & Judge, T.A. (2001). Personality and performance at the beginning of the new millennium: what do we know and where do we go next? *International Journal of Selection and Assessment*, 9(1), 9-30.
- Becker, G. (1976). *The economic approach to human behavior*. Chicago: The Chicago University Press.
- Bennis, W. (2003). *On becoming a leader*. Perseus Publishing.
- Berry, D.S. (2000). Attractiveness, attraction, and sexual selection: Evolutionary perspectives on the form and function of physical attractiveness. In P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 273-342). San Diego, CA: Academic Press.
- Betsch, T., Brinkmann, B.J., Fiedler, K. & Breinig, K. (1999). When prior knowledge overrules new evidences: Adaptive use of decision strategies and the role of behavioral routines. *Swiss Journal of Psychology*, 58(3), 151-160
- Bierhoff, H.-W. (2000). *Sozialpsychologie. Ein Lehrbuch*. Stuttgart, Berlin, Köln: Kohlhammer Verlag.
- Biernat, M. (1993). Gender and height: Developmental patterns in knowledge and use of an accurate stereotype. *Sex Roles*, 29, 691-713.
- Bligh, M.C. & Meindl, J.R. (2005). The cultural ecology of leadership: An analysis of popular leadership books. In D.M. Messick & R.M. Kramer (Eds.), *The psychology of leadership. New perspectives and research* (pp.11-52). Mahwah, NJ: Lawrence Erlbaum.
- Boehm, C. (1999). *Hierarchy in the forest*. London: Harvard University Press.
- Bono, J.E. & Judge, T.A. (2004). Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology*, 89, 901-910.
- Borkenau, P. & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology*, 65, 546-553.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F.M. & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intel-

- ligence. *Journal of Personality and Social Psychology*, 86(4), 599-614.
- Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon Press.
- Bröder, A. (2003). Decision making with the >adaptive toolbox<: Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Social Psychology*, 29, 611-625.
- Bruner, J.S. & Postman, L. (1951). An approach to social perception. In W. Dennis & R. Lippitt (Eds.), *Current trends in social psychology* (pp. 71-118). Pittsburgh: University of Pittsburgh Press.
- Buchanan, W. & Cantril, H. (1953). *How nations see each other*. Urbana: University of Illinois Press.
- Budesheim, T. L. & De Paola, A. J. (1994). Beauty or beast? The effects of appearance, personality, and issue information on evaluations of political candidates. *Personality and Social Psychology Bulletin*, 20, 339- 348
- Buss, D.M. (2004). *Evolutionäre Psychologie*. München: Pearson
- Carney, D.R., Colvin, C.R. & Hall, J.A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, 41, 1054-1072.
- Christiansen, N.D., Wolcott-Burnam, S., Janovics, J.E., Burns, G.N. & Quirk, S.W. (2005). The good judge revisited: Individual differences in the accuracy of personality judgments. *Human Performance*, 18(2), 123-149.
- Cloutier, J., Mason, M.F. & Macrae, C.N. (2005). The perceptual determinants of person construal: Reopening the social-cognitive toolbox. *Journal of Personality and Social Psychology*, 88(6), 885-894.
- Collins, A.F. (1999). The enduring appeal of physiognomy: Physical appearance as a sign of temperament, character, and intelligence. *History of Psychology*, 2(4), 251-276.
- Conelly, M.S., Gilbert, J.A., Zaccaro, S.J., Threfall, K.V., Marks, M.A. & Mumford, M.D. (2000). Exploring the relationship of leadership skills and knowledge to job performance. *Leadership Quarterly*, 11, 65-86.
- Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163-228). Oxford: Oxford University Press.

- Costa Jr., P.T., Terracciano, A. & McCrae, R.R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81(2), 322-331.
- Crant, A.M. & Bateman, T.S. (2000). Charismatic leadership viewed from above: the impact of proactive personality. *Journal of Organizational Behavior*, 21, 63-75.
- Cronbach, L.J. (1955). Processes affecting scores on „understanding of others“ and „assumed similarity“. *Psychological Bulletin*, 52, 177-193.
- DePaulo, B.M., Blank, A.L., Swaim, G.W. & Hairfield, J.G. (1992). Expressiveness and expressive control. *Personality and Social Psychology Bulletin*, 18(3), 276-285.
- Deaux, K. & LaFrance, M. (1998). Gender. In D.T. Gilovich, S.T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 788-827). Boston, Mass: McGraw-Hill.
- Diekmann, A.B. & Eagly, A.H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin*, 26(10), 1171-1188.
- Diekmann, A.B., Eagly, A.H., Mladinic, A. & Ferreira, M.C. (2005). Dynamic stereotypes about women and men in Latin America and the United States. *Journal of Cross-Cultural Psychology*, 36(2), 209-226.
- Dion, K., Berscheid, E. & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285-290.
- Dymond, R.F. (1950). Personality and Empathy. *Journal of Consulting Psychology*, 14, 343-350.
- Eagly, A. H. (2003). The rise of female leaders. *Zeitschrift für Sozialpsychologie*, 34, 123-132.
- Eagly, A. H. & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573-598.
- Eagly, A.H., Karau, S.J. & Makhijani, M.G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, 117(1), 125-145.
- Eagly, A. H., Makhijani, M. G. & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis, *Psychological Bulletin*, 111, 3-22.
- Eagly, A.H. & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender ste-

- reotypes, and judgements of competence. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 5, pp. 1-35). Chichester, UK: Wiley.
- Eagly, A. H., & Sczesny, S. (in press). Stereotypes about women, men, and leaders: Have times changed? In M. Barreto, M. Ryan, & M. Schmitt (Eds.), *Barriers to diversity: The glass ceiling after 20 years*. Washington, DC: APA Books.
- Eagly, A.H. & Steffen, J.F. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46, 735-754.
- Ebner, B., Gathman, S. & Wiedermann, A. (2002). Schönheit und der Halo-Effekt. In A. Hergovich (Hrsg.), *Psychologie der Schönheit. Physische Attraktivität aus wissenschaftlicher Perspektive* (S. 187-204). Wien: WUV-Universitätsverlag.
- Efran, M. G. & Patterson, E. W. J. (1974). Voters vote beautiful: The effect of physical appearance on a national election. *Canadian Journal of Behavioral Science*, 6(4), 352-356.
- Eilles-Mathiessen, C., el Hage, N., Janssen, S. & Osterholz, A. (2002). *Schlüsselqualifikationen in Personalauswahl und Personalentwicklung*. Bern: Verlag Hans Huber.
- Ettcoff, N. (1999). *Survival of the prettiest*. London: Abacus
- Evans, J.St.B.T. (1984). In defense of the citation bias in the judgment literature [Comment]. *American Psychologist*, 39, 1500-1501.
- Feingold, A. (1992). Good looking people are not what we think. *Psychological Bulletin*, 111(2), 304-341.
- Fetchenhauer, D., Groothuis, T. & Pradel, J. (under review). Not only states but traits – Humans can identify permanent altruistic dispositions in 20 seconds. *Manuscript submitted for publication in Evolution and Human Behavior*.
- Fiedler, F.E. (1967). *A theory of leadership effectiveness*. New York: McGraw Hill.
- Fiedler, F.E. & Macaulay, J.L. (1998). The leadership situation: A missing factor in selecting and training managers. *Human Resource Management Review*, 8(4), 335-350.
- Fischer, L. & Wiswede, G. (2002). *Grundlagen der Sozialpsychologie*. München, Wien: R. Oldenbourg Verlag.
- Fisek, M.H. & Ofshe, R. (1970). The process of status evolution. *Sociometry*, 33, 327-346.

- Fleenor, J.W., Taylor, S. & Chappelow, C. (2008). *Leveraging the impact of 360-degree feedback*. San Francisco: Wiley.
- Försterling, F., Preikschas, S. & Agthe, M. (2007). Ability, luck, and looks: An evolutionary look at achievement ascriptions and the sexual attribution bias. *Journal of Personality and Social Psychology*, 92(5), 775-788.
- Fox, E. & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotions*, 6(3), 534-539.
- Fox, E., Griggs, L. & Mouchlianitis, E. (2007). The detection of fear-relevant stimuli: Are guns noticed as quickly as snakes? *Emotion*, 7(4), 691-696.
- Freire, A., Eskritt, M. & Lee, K. (2004). Are eyes windows to a deceiver's soul? Children's use of another's eye gaze cues in a deceptive situation. *Developmental Psychology*, 40, 1093-1104.
- Funder, D.C. (1980). The „trait“ of ascribing traits: Individual differences in the tendency to trait ascription. *Journal of Research in Personality*, 14, 376-385.
- Funder, D.C. (1987). Errors and Mistakes: Evaluating the Accuracy of Social Judgment. *Psychological Bulletin*, 101(1), 75-90.
- Funder, D.C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652-670.
- Funder, D.C. (1999). *Personality Judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.
- Gadagkar, R. (2003). Is the peacock merely beautiful or also honest? *Current Science*, 85, 1012-1020.
- Gergen, K.J. (2001). *Social construction in context*. London, Thousand Oaks, New Delhi: Sage Publications.
- Gigerenzer, G. (2007). *Bauchentscheidungen*. München: C. Bertelsmann Verlag.
- Gigerenzer, G. (2002a). *Adaptive thinking. Rationality in the real world*. New York, NY: Oxford University Press.
- Gigerenzer, G. (2002b). *Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken*. Berlin: Berlin Verlag.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, 103(3), 592-596.
- Gigerenzer, G. (1982). Der eindimensionale Wähler. *Zeitschrift für Sozialpsychologie*, 13, 284-313.

- Gigerenzer, G. & Goldstein, D.G. (2001). Betting on one good reason: The Take The Best heuristic. In G. Gigerenzer, P.M. Todd & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75-95), New York, NY: Oxford University Press.
- Gigerenzer, G., Todd, P.M. & the ABC Research Group (2001). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.
- Gillen, B. (1981). Physical attractiveness: A determinant of two types of goodness. *Personality & Social Psychology Bulletin*, 7, 277-281.
- Gilovich, T. & Griffin, D. (2002). Introduction – Heuristics and Biases: Then and now. In T. Gilovich, D. Griffin & D. Kahnemann (Eds.), *Heuristics and biases: the psychology of intuitive judgment* (pp. 1-18), New York, NY: Cambridge University Press.
- Gilovich, T., Griffin, D. & Kahnemann, D. (2002). *Heuristics and biases: the psychology of intuitive judgment*. New York, NY: Cambridge University Press.
- Gilovich, T., Keltner, D. & Nisbett, R.E. (2006). *Social psychology*. New York, London: W. W. Norton & Company, Inc.
- Hannover, B. & Kessels, U. (2003). Erklärungsmuster weiblicher und männlicher Spitzen-Manager zur Unterrepräsentanz von Frauen in Führungspositionen. *Zeitschrift für Sozialpsychologie*, 34, 197-204.
- Harackiewicz, J.M. & DePaulo, B.M. (1982). Accuracy of person perception: A component analysis according to Cronbach. *Personality and Social Psychology Bulletin*, 8, 247-256.
- Hartley, L (2001). *Physiognomy and the meaning of expression in the nineteenth-century culture*. New York, NY: Cambridge University Press.
- Haselton, M.G. & Funder, D.C. (2006). The evolution of accuracy and bias in social judgement. In M. Schaller, J.A. Simpson & D.T. Kenrick (Eds.), *Evolution and Social Psychology* (pp. 15-37). New York, Hove: Psychology Press.
- Heilman, M.E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, 5, 269-298.
- Heilman, M.E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57(4), 657-674.
- Heilman, M.E. & Saruwatari, L.R. (1979). When beauty is beastly: The effects of appearance and sex on evaluations of job appli-

- cants for managerial and nonmanagerial jobs. *Organizational Behavior and Human Performance*, 23, 360-372.
- Heilman, M.E. & Stopeck, M.H. (1985). Being attractive, advantage or disadvantage? Performance-based evaluations and recommended personnel actions as a function of appearance, sex, and job type. *Organizational Behavior and Human Decision Processes*, 35, 202-215.
- Heilmann, K. (2002). Die Konstruktionsübung: Eine besondere Übung im Assessment-Center. In E. Fay (Hg.), *Das Assessment-Center in der Praxis. Konzepte-Erfahrungen-Innovationen* (S. 103-129). Göttingen: Vandenhoeck & Ruprecht.
- Hewstone, M. (1994). Revising and change of stereotypic beliefs: In search of the elusive subtyping model. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 5, pp. 69-109). New York: Wiley.
- Hodapp, V. & Rohrmann, S. (2006). Differentielle Psychologie: Persönlichkeitsforschung. In K. Pawlik (Hrsg.), *Handbuch Psychologie. Wissenschaft-Anwendung-Berufsfelder* (S.337-353). Heidelberg: Springer Medizin Verlag.
- Höft, S. & Funke, U. (2006). Simulationsorientierte Verfahren in der Personalauswahl. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S.145-188). Göttingen: Hogrefe Verlag.
- Hogarth, R.M. (2001). *Educating intuition*. Chicago: The University Press.
- Hogarth, R.M. (2005). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. In T. Betsch & S. Haberstroh (Eds.), *The routines of decision making* (pp. 67-82). Mahwah, NJ: Lawrence Erlbaum Associates.
- House, R.J. & Howell, J.M. (1992). Personality and charismatic leadership. *Leadership Quarterly*, 3, 81-108.
- Hussy, W. (1986). *Denkpsychologie. Bd. 2. Denken, Urteilen, Kreativität, Sprache, Entwicklung*. Stuttgart, Berlin, Köln, Mainz: Kohlhammer.
- Jackson, D., Engstrom, E. & Emmers-Sommer, T. (2007). Think leader, think male and female: Sex vs. Seating arrangement as leadership cues. *Sex Roles*, 57, 713-723.
- Jacob, S. (2004). *Soziale Repräsentationen und Relationale Realitäten*. Wiesbaden: Deutscher Universitäts-Verlag/GWV Fachverlage GmbH.

- Johnson, A.W. & Earle, T. (2000). *The evolution of human societies*. Stanford, CA: Stanford University Press.
- Jones, E.E. & Harris, V.A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, 3, 1-24.
- Judd, C.M. & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100(1), 109-128.
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Kahneman, D. & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103(3), 582-591.
- Kalma, A. (1991). Hierarchisation and dominance assessment at first glance. *European Journal of Social Psychology*, 21, 165-181.
- Kammrath, L.K., Mendoza-Denton, R. & Mischel, W. (2005). Incorporating if... then... personality signatures in person perception: Beyond the person-situation dichotomy. *Journal of Personality and Social Psychology*, 88(4), 605-618.
- Kanning, U.P. (1999). *Die Psychologie der Personenbeurteilung*. Göttingen, Bern, Toronto, Seattle: Hogrefe.
- Kebeck, G. (1997). *Wahrnehmung. Theorien, Methoden und Forschungsergebnisse der Wahrnehmungspsychologie*. Weinheim, München: Juventa Verlag.
- Kelley, H.H. (1979). *Personal relationships: Their structure and processes*. Hillsdale, NJ: Erlbaum.
- Kenny, D.A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, 98, 155-163.
- Kenny, D.A., Albright, L., Malloy, T.E. & Kashy, D. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin*, 116, 245-258.
- Kenrick, D.T., Schaller, M. & Simpson, J.A. (2006). Evolution ist the new cognition. In M. Schaller, J.A. Simpson & D.T. Kenrick (Eds.), *Evolution and Social Psychology* (pp. 1-13). New York, Hove: Psychology Press.
- Klein, G. (1998). *Sources of power: How people make decisions*. Massachusetts: MIT Press.
- Klein, G. (2004). *The power of intuition. How to use your gut feelings to make better decisions at work*. New York: Doubleday.
- Kleinmann, M. (2003). *Assessment Center*. Göttingen: Hogrefe

- Krueger, J.I. & Funder, D.C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral And Brain Sciences*, 27, 313-376.
- Kunda, Z. (1999). *Social cognition. Making sense of people*. Cambridge, London: The MIT Press.
- Kunda, Z. & Oleson, K.C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping. *Journal of Personality and Social Psychology*, 68, 565-579.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M. & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390- 423.
- Lievens, F. & Thornton III, G.C. (2005). Assessment Centers : Recent developments in practice and research. In A. Evers, N. Anderson & O. Voskuil (Eds.), *The Blackwell Handbook of Personnel Selection* (pp. 243-264). Malden, Oxford, Carlton: Blackwell Publishing.
- Lippa, R. & Dietz, J.K. (2000). The relation of gender, personality, and intelligence to judges' accuracy in judging strangers' personality from brief video segments. *Journal of Nonverbal Behavior*, 24, 25-43.
- Locher, P., Unger, R., Sociedade, P. & Wahl, J. (1993). At first glance: Accessibility of the physical attractiveness stereotype. *Sex Roles*, 28, 729-743.
- Locksley, A., Borgida, E., Brekke, N. & Hepburn, C. (1980). Sex stereotypes and social judgement. *Journal of Personality and Social Psychology*, 39(5), 821-831.
- Locksley, A., Hepburn, C. & Ortiz, V. (1982). Social stereotypes and judgements of individuals : An instance of the base-rate fallacy. *Journal of Experimental Social Psychology*, 18, 23-42.
- Lord, R.G., DeVader, C.L. & Alliger, G.M. (1986). A meta analysis of the relation between personality traits and leadership perceptions : An application of validity generalization procedures. *Journal of Applied Psychology*, 71, 402-410.
- Luxen, M.F. & Van De Vijver, F.J.R. (2006). Facial attractiveness, sexual selection, and personnel selection: When evolved preferences matter. *Journal of Organizational Behavior*, 27, 241-255.
- Maner, J.K., Kenrick, D.T., Becker, D.V., Delton, A.W., Hofer, B., Wilbur, C.J., et al. (2003). Sexually selective cognition: Beauty

- captures the mind of the beholder. *Journal of Personality and Social Psychology*, 85(6), 1107-1120.
- Mann, R.D. (1959). A review of the relationship between personality and performance in small groups. *Psychological Bulletin*, 56, 214-270.
- Marcus, B. & Schuler, H. (2006). Leistungsbeurteilung. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S.145-188). Göttingen: Hogrefe Verlag.
- Marlow, C.M., Schneider, S.L. & Nelson, C.E. (1996). Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased? *Journal of Applied Psychology*, 81(1), 11-21.
- Merton, R.K. (1948). The self-fulfilling prophecy. *Antioch Review*, 8, 193-210.
- Murphy, N.A., Hall, J.A. & Colvin, R.C. (2003). Accurate intelligence assessments in social interactions: mediators and gender effects. *Journal of Personality*, 71, 465-493.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, 55, 1-22.
- Nisbett, R. & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement*. New York: Prentice Hall.
- Noelle-Neumann, E. (1983). *Spiegel-Dokumentation: Persönlichkeitsstärke. Ein neuer Maßstab zur Bestimmung von Zielgruppenpotentialen*. Hamburg: Spiegel-Verlag.
- Ostroff, C. & Judge, T.A. (2007). *Perspectives on organizational fit*. New York, NY: Lawrence Erlbaum Associates.
- Öhman, A. & Mineka, S. (2001). Fear, phobia, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483-522.
- Offermann, L.R., Kennedy Jr., J.K. & Wirt, P.W. (1994). Implicit leadership theories: Content, structure and generalizability. *Leadership Quarterly*, 5, 43-58.
- Pawlik, K. (2006). *Handbuch Psychologie*. Heidelberg: Springer Verlag.
- Penton-Voak, I.S., Pound, N., Little, A.C. & Perrett, D.I. (2006). Personality judgements from natural and composite facial images: more evidence for a „kernel of truth“ in social perception. *Social Cognition*, 24(5), 607-640.

- Petrie, M. & Halliday, T. (1994). Experimental and natural changes in the peacock's (*Pavo cristatus*) train can affect mating success. *Behavioral and Ecological Sociobiology*, 35, 213-217.
- Pettigrew, T.F. (1964). *A profile of the negro american*. Washington: Howard University Press
- Plessner, H., Betsch, C. & Betsch, T. (2007). *Intuition in judgment and decision making*. Mahwah, NJ: Lawrence Erlbaum Associates
- Plessner, H. & Czenna, S. (2007). The benefits of intuition. In H. Plessner, C. Betsch & T. Betsch (Eds.), *Intuition in judgement and decision making*. NY: Manwah.
- Powell, G. N., Butterfield D. A. & Parent, J. D. (2002). Gender and managerial stereotypes: Have the times changed? *Journal of Management*, 28, 177-193.
- Precht, R.D. (2007). *Wer bin ich und wenn ja, wie viele?* München: Wilhelm Goldmann Verlag.
- Prothro, E.T. & Melikian, .H. (1955). Studies in stereotypes: V. Familiarity and the kernel of truth hypothesis. *The Journal of Social Psychology*, 41, 3-10.
- Rao, T.V. & Rao, R (2005). *The power of 360 degree feedback: maximizing mangerial and leadership effectiveness*. New Delhi: Rspnse Books.
- Reynolds, D.J. & Gifford, R. (2001). The sounds and sights of intelligence: A lens model channel analysis. *Personality and Social Psychology Bulletin*, 27, 187-200.
- Richerson, P. J., & Boyd, R. (2001). Built for speed, not for comfort: Darwinian theory and human culture. *History and Philosophy of the Life Sciences*, 23, 423-463.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone*. Chicago, IL: University of Chicago Press.
- Roszell, P., Kennedy, D. & Grabb, E. (2001). Physical attractiveness and income among canadians. *The Journal of Psychology*, 123(6), 547-559.
- Schaller, M. (2008). Evolutionary bases of first impressions. In N. Ambady & J.J. Skowronski (Eds.), *First Impressions*. New York: Guilford Press.
- Schaller, M., Park, J.H. & Kenrick, D.T. (2007). Human evolution and social cognition. In R.I.M. Dunbar & L. Barrett (Eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 491-504). Oxford University Press.

- Schein, V. E. (1973). The relation between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, 57, 95-100.
- Schein, V. E. (1975). Relations between sex-role-stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology*, 60, 340-344.
- Schein, V. E. (2001). A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, 57(4), 675-688.
- Schein, V. E., Mueller, R., Lituchy, T. & Liu, J. (1996). Think manager – think male: A global phenomenon? *Journal of Organizational Behavior*, 17, 33-41.
- Scherm, M. & Kaufel, S. (2005). 360-Grad-Feedback. In I. Jöns & W-Bungard (Hrsg.), *Feedbackinstrumente im Unternehmen* (S.113-128). Wiesbaden: Gabler Verlag.
- Schmidt, F.L. & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86(1), 162-173.
- Schuler, H. (2006). Stand und Perspektiven der Personalpsychologie. *Zeitschrift für Arbeits- und Organisationspsychologie*, 4, 176-188.
- Seidel, E. (1978). *Betriebliche Führungsformen*. Stuttgart: Poeschel.
- Shamir, B., House, R.J. & Arthur, M.B. (1993). The motivational effects of charismatic leadership: A self-concept based theory. *Organization Science*, 4, 167-177.
- Shamir, B., Zakay, E., & Breinin, E. (1998). Correlates of charismatic leader behavior in military units: Subordinates'attitudes, unit characteristics, and superiors'appraisals of leader performance. *Academy of Management Journal*, 41, 387-409.
- Simon, H.A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 99-118.
- Sczesny, S. (2003a). A Closer Look Beneath the Surface: Various Facets of the Think-Manager – Think-Male Stereotype. *Sex Roles*, 49, 353-363.
- Sczesny, S. (2003b). Führungskompetenz: Selbst- und Fremdwahrnehmung weiblicher und männlicher Führungskräfte. *Zeitschrift für Sozialpsychologie*, 34, 133-145.
- Sczesny, S., Bosak, J., Neff, D. & Schyns, B. (2004). Gender stereotypes and the attribution of leadership traits: A cross-cultural comparison. *Sex Roles*, 51, 11/12, 631- 645.

- Sczesny, S. & Kühnen, U. (2004). Sex and gender-stereotypic physical appearance: Consequences for the assessment of leadership competence. *Personality and Social Psychology Bulletin*, 30, 13–21.
- Sczesny, S., Spreemann, S. & Stahlberg, D. (2006). Masculine = Competent? Physical Appearance and Sex as Sources of Gender-Stereotypic Attributions. *Swiss Journal of Psychology*, 65, 15–23.
- Sieverding, M. (2003). Frauen unterschätzen sich: Selbstbeurteilungs-Biases in einer simulierten *Bewerbungssituation*. *Zeitschrift für Sozialpsychologie*, 34(3), 147–160.
- Slater, A., Von der Schulenberg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S. et al. (1998). Newborn infants prefer attractive faces. *Infant Behavior & Development*, 21, 345–354.
- Snyder, M., Decker Tanke, E. & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 35, 656–666.
- Steiger, J.H. & Browne, M.W. (1984). The comparison of interdependent correlations between optimal linear composites. *Psychometrika*, 49(1), 11–24.
- Stogdill, R.M. (1948). Personal factors associated with leadership. *Journal of Psychology*, 25, 35–71.
- Streukens, S. & De Ruyter, K. (2004). Reconsidering nonlinearity and asymmetry in customer satisfaction and loyalty models: An empirical study in three retail service settings. *Marketing Letters*, 15, 99–111.
- Sugiyama, L.S. (2005). Physical attractiveness in adaptionist perspective. In D.M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (pp. 292 – 343). Hoboken, NJ: John Wiley & Sons, Inc.
- Surowiecki, J. (2005). *Die Weisheit der Vielen*. München: Wilhelm Goldmann Verlag.
- Swazina, K.R., Waldherr, K. & Maier, K. (2004). Geschlechtsspezifische Ideale im Wandel der Zeit. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25(3), 165–176.
- Taft, R. (1955). The ability to judge people. *Psychological Bulletin*, 52, 1–23.
- Todd, P.M. & Gigerenzer, G. (2003). Bounding rationality to the world. *Journal of Economic Psychology*, 24, 143–165.

- Van Vugt, M., Hogan, R. & Kaiser, R.B. (2008). Leadership, followership, and evolution. Some lessons from the past. *American Psychologist*, 63(3), 182-196.
- Vogt, D.S. & Colvin, C.R. (2003). Interpersonal orientation and the accuracy of personality judgments. *Journal of Personality*, 71, 267-295.
- Von Rosenstiel, L. (2006). Führung. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S.353-384). Göttingen: Hogrefe Verlag.
- Vroom, V.H. (2000). Leadership and decision-making process. *Organizational Dynamics*, 28, 82-94
- Vroom, V.H. & Yetton, P. (1973). *Leadership and decision-making*. Pittsburgh: University of Pittsburgh Press.
- Wänke, M., Bless, H. & Wortberg, S. (2003). Der Einfluss von Karrierfrauen auf das Frauenstereotyp. Die Auswirkungen von Inklusion und Exklusion. *Zeitschrift für Sozialpsychologie*, 34, 187-196.
- Waxer, P.H. (1976). Nonverbal cues for depth of depression: Set versus no set. *Journal of Consulting and Clinical Psychology*, 44, 93.
- Waxer, P.H. (1977). Nonverbal cues for anxiety: An examination of emotional leakage. *Journal of Abnormal Psychology*, 86, 306-314.
- Weber, M. (1947). *The theory of social and economic organizations*. New York: Free Press.
- Weber, R. & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45, 961-977.
- Weimann, G. (1991). The influentials: Back to the concept of opinion leaders? *Public Opinion Quarterly*, 55, 267-297.
- Weiss, W.H. (2003). Effective leadership: what are the requisites? *Supervision*, 61, 3-6.
- Wellington, S., Kropf, M.B. & Gerkovich, P.R. (2003). What's holding women back? *Harvard Business Review*, 81(6), 18-19.
- Yukl, G. (2002). *Leadership in Organizations*. New Jersey: Prentice Hall.
- Zebrowitz, L.A., Hall, J.A., Murphy, N.A. & Rhodes, G. (2002). Looking smart and looking good: facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, 28, 238-249.

9. Anhang

Anhang A: Skala Persönlichkeitsstärke (Noelle-Neumann, 1983):

	Trifft zu	Trifft nicht zu
1. Gewöhnlich rechne ich bei dem, was ich mache, mit Erfolg.	<input type="checkbox"/>	<input type="checkbox"/>
2. Ich bin selten unsicher, wie ich mich verhalten soll.	<input type="checkbox"/>	<input type="checkbox"/>
3. Ich übernehme gerne Verantwortung.	<input type="checkbox"/>	<input type="checkbox"/>
4. Ich übernehme bei gemeinsamen Unternehmungen gern die Führung.	<input type="checkbox"/>	<input type="checkbox"/>
5. Es macht mir Spaß, andere Menschen von meiner Meinung zu überzeugen.	<input type="checkbox"/>	<input type="checkbox"/>
6. Ich merke öfter, dass sich andere nach mir richten.	<input type="checkbox"/>	<input type="checkbox"/>
7. Ich kann mich gut durchsetzen.	<input type="checkbox"/>	<input type="checkbox"/>
8. Ich bin anderen oft einen Schritt voraus.	<input type="checkbox"/>	<input type="checkbox"/>
9. Ich besitze vieles, worum mich andere beneiden.	<input type="checkbox"/>	<input type="checkbox"/>
10. Ich gebe anderen Ratschläge/Empfehlungen.	<input type="checkbox"/>	<input type="checkbox"/>
11. Berufliche Führungsposition, Stellung als Vorgesetzter?	<input type="checkbox"/>	<input type="checkbox"/>
12. Während der Freizeit Mitarbeit in einer Partei, Gewerkschaft, Bürgerinitiative?	<input type="checkbox"/>	<input type="checkbox"/>
13. Ausübung eines Amtes in einem Verein, Organisation, etc.?	<input type="checkbox"/>	<input type="checkbox"/>

Anhang B: Beurteiler-Instruktionen für die Einschätzung der Persönlichkeitsstärke der Zielpersonen

Liebe Studierende,

herzlich Willkommen zu diesem Experiment. Vielen Dank, dass Sie sich bereit erklärt haben, an dieser Studie teilzunehmen.

Diese Untersuchung erfolgt vollständig anonym.

Im Folgenden werden Sie einige Videoclips sehen. Jeder Videoclip zeigt eine Person. Zwischen den einzelnen Szenen, werden Sie ca. 10 Sekunden Zeit haben, um folgende drei Fragen zu der jeweils im letzten Clip vorgestellten Person zu beantworten. Bitte kreuzen Sie jeweils an, in welchem Ausmaß die Aussage ihrer Meinung nach auf die Person zutrifft.

Die Antwortkategorien gehen von

1 („trifft gar nicht zu“) bis 7 („trifft voll und ganz zu“).

Person 1...	Trifft gar nicht zu						Trifft voll und ganz zu
... rechnet bei dem, was sie macht, gewöhnlich mit Erfolg.	1	2	3	4	5	6	7
... übernimmt bei gemeinsamen Unternehmungen gerne die Führung.	1	2	3	4	5	6	7
... kann sich gut durchsetzen.	1	2	3	4	5	6	7

Dieser Versuch ist eine Studie zum ersten Eindruck. Es gibt daher kein richtig oder falsch. Uns interessiert, wie Sie diese Person nach Ihrem ersten Eindruck einschätzen

Vorab schon einmal herzlichen Dank für Ihre Mitarbeit an dieser Untersuchung!

Anhang C: Beurteiler-Instruktionen für die Feststellung der Attraktivität der Zielpersonen

Liebe Studierende,

herzlich Willkommen zu diesem Experiment. Vielen Dank, dass Sie sich bereit erklärt haben, an dieser Studie teilzunehmen.

Im Folgenden werden Sie einige Videoclips sehen. Jeder Videoclip zeigt eine Person. Zwischen den einzelnen Szenen, werden Sie ca. 10 Sekunden Zeit haben, um folgende Frage zu der jeweils im letzten Clip vorgestellten Person zu beantworten. Bitte kreuzen Sie jeweils an, in welchem Ausmaß die Aussage Ihrer Meinung nach auf die Person zutrifft.

Person 1 finde ich...								
überhaupt nicht attraktiv	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	sehr attraktiv

Dieser Versuch ist eine Studie zum ersten Eindruck. Es gibt daher kein richtig oder falsch. Uns interessiert, wie attraktiv Sie diese Person nach dem ersten Eindruck finden.

Vorab schon einmal herzlichen Dank für Ihre Mitarbeit an dieser Untersuchung!

Anhang D: Beobachtungsbogen der Verhaltensauszählung in Studie 2

Fragebogennummer: <div style="border-bottom: 1px solid black; width: 100px; margin-top: 5px;"></div>	Beobachter: <div style="border-bottom: 1px solid black; width: 100px; margin-top: 5px;"></div>			
Einstieg finden/ Aufgabe beginnen:				
Hand am Werkzeug:				
Vorschläge / Lösungsansätze:				
Lautstärke:	<table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="width: 30px; height: 30px; vertical-align: middle;">-</td> <td style="width: 30px; height: 30px; vertical-align: middle; text-align: center;">0</td> <td style="width: 30px; height: 30px; vertical-align: middle; text-align: right;">+</td> </tr> </table>	-	0	+
-	0	+		
Entscheidungen treffen:				
Aufgaben verteilen:				
Moderationsfunktionen:				
<ul style="list-style-type: none"> • Uhrzeit 				
<ul style="list-style-type: none"> • Ziel vor Augen 				
<ul style="list-style-type: none"> • alle einbeziehen / motivieren 				

Anhang E: Fragebogen zur Erhebung des 360°-Feedbacks in Studie 4

Liebe/r Studienteilnehmer/in!

Vielen Dank, dass Sie bereit sind, den vorliegenden Fragebogen auszufüllen.

Im Folgenden werden Sie gebeten, die Person einzuschätzen, die Ihnen diesen Umschlag ausgehändigt hat. Dabei gibt es weder richtige und falsche noch gute und schlechte Antworten. Es geht lediglich um Ihre persönliche Einschätzung. Diese bleibt völlig anonym, weshalb Sie auch an keiner Stelle Ihren Namen nennen.

Bitte lesen Sie sich die Instruktionen sorgfältig durch und beantworten Sie die anschließenden zügig und aus dem Bauch heraus.

Wie lange kennen Sie die einzuschätzende Person:

Die einzuschätzende Person ist mein/e:

Ihr Geschlecht: ☐ männlich ☐ weiblich

Bitte kreuzen Sie nun an, in welchem Ausmaß die folgenden Aussagen auf die zu beurteilende Person zutreffen. Die Antwortkategorien gehen von 1 („trifft gar nicht zu“) bis 7 („trifft voll und ganz zu“).

Die Person, die ich einschätze...							
... rechnet bei dem, was sie macht, gewöhnlich mit Erfolg.	1	2	3	4	5	6	7
... übernimmt bei gemeinsamen Unternehmungen gerne die Führung.	1	2	3	4	5	6	7
... kann sich gut durchsetzen.	1	2	3	4	5	6	7

Vielen Dank für Ihre Teilnahme. Sie haben uns sehr geholfen!

Bitte stecken Sie den ausgefüllten Fragebogen in den beigelegten Umschlag und geben ihn in die Post.